GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:18:01 ; Search time 35.15 Seconds
(without alignments)
22.237 Million cell updates/sec

Title: US-09-461-061A-4
Perfect score: 161
Sequence: 1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 161 | 100.0 | 117 | 1 | US-08-193-114B-1 | Sequence 1, Appli |
| 2 | 161 | 100.0 | 117 | 5 | PCT-US92-06809-1 | Sequence 1, Appli |
| 3 | 95 | 59.0 | 26 | 4 | US-08-676-242-15 | Sequence 15, Appl |
| 4 | 52 | 32.3 | 701 | 2 | US-08-533-669A-16 | Sequence 16, Appl |
| 5 | 49.5 | 30.7 | 352 | 2 | US-08-483-926A-11 | Sequence 11, Appl |
| 6 | 49.5 | 30.7 | 352 | 2 | US-08-737-045-12 | Sequence 12, Appl |
| 7 | 48 | 29.8 | 145 | 2 | US-08-832-535-2 | Sequence 2, Appli |
| 8 | 48 | 29.8 | 145 | 3 | US-09-019-485-2 | Sequence 2, Appli |
| 9 | 48 | 29.8 | 145 | 3 | US-09-019-485-3 | Sequence 3, Appli |
| 10 | 48 | 29.8 | 145 | 4 | US-09-431-480-9 | Sequence 9, Appli |
| 11 | 48 | 29.8 | 145 | 4 | US-09-617-302-9 | Sequence 9, Appli |
| 12 | 48 | 29.8 | 178 | 2 | US-08-791-522-1 | Sequence 1, Appli |
| 13 | 48 | 29.8 | 178 | 3 | US-09-314-777-1 | Sequence 1, Appli |
| 14 | 48 | 29.8 | 350 | 1 | US-08-415-751-20 | Sequence 20, Appl |
| 15 | 46 | 28.6 | 126 | 4 | US-08-751-359-11 | Sequence 11, Appl |
| 16 | 46 | 28.6 | 126 | 4 | US-08-907-146-11 | Sequence 11, Appl |
| 17 | 46 | 28.6 | 128 | 1 | US-08-259-372A-14 | Sequence 14, Appl |
| 18 | 46 | 28.6 | 128 | 1 | US-08-468-671-14 | Sequence 14, Appl |
| 19 | 46 | 28.6 | 1220 | 2 | US-08-843-530B-36 | Sequence 36, Appl |
| 20 | 45.5 | 28.3 | 671 | 6 | 5266464-2 | Patent No. 5266464 |
| 21 | 45 | 28.0 | 108 | 2 | US-08-378-939-32 | Sequence 32, Appl |
| 22 | 45 | 28.0 | 108 | 2 | US-08-378-939-34 | Sequence 34, Appl |
| 23 | 45 | 28.0 | 148 | 5 | PCT-US95-07135-2 | Sequence 2, Appli |
| 24 | 45 | 28.0 | 149 | 2 | US-08-461-030C-2 | Sequence 2, Appli |
| 25 | 45 | 28.0 | 149 | 3 | US-08-744-138-2 | Sequence 2, Appli |
| 26 | 45 | 28.0 | 149 | 4 | US-09-431-480-8 | Sequence 8, Appli |
| 27 | 45 | 28.0 | 149 | 4 | US-09-431-480-10 | Sequence 10, Appl |
| 28 | 45 | 28.0 | 149 | 4 | US-09-617-302-8 | Sequence 8, Appli |
| 29 | 45 | 28.0 | 149 | 4 | US-09-617-302-10 | Sequence 10, Appl |
| 30 | 45 | 28.0 | 149 | 4 | US-09-241-376-2 | Sequence 2, Appli |
| 31 | 45 | 28.0 | 620 | 4 | US-09-269-731-6 | Sequence 6, Appli |
| 32 | 44.5 | 27.6 | 872 | 1 | US-08-766-014-2 | Sequence 2, Appli |
| 33 | 44 | 27.3 | 252 | 1 | US-08-460-512-5 | Sequence 5, Appli |
| 34 | 44 | 27.3 | 285 | 1 | US-08-460-512-2 | Sequence 2, Appli |
| 35 | 43.5 | 27.0 | 415 | 4 | US-08-675-816-2 | Sequence 2, Appli |
| 36 | 43.5 | 27.0 | 532 | 3 | US-08-481-435-12 | Sequence 12, Appl |
| 37 | 43.5 | 27.0 | 553 | 3 | US-08-481-435-11 | Sequence 11, Appl |
| 38 | 43.5 | 27.0 | 823 | 3 | US-08-481-435-4 | Sequence 4, Appli |
| 39 | 43.5 | 27.0 | 836 | 3 | US-08-481-435-9 | Sequence 9, Appli |
| 40 | 43.5 | 27.0 | 844 | 3 | US-08-481-435-7 | Sequence 7, Appli |
| 41 | 43.5 | 27.0 | 844 | 3 | US-08-481-435-8 | Sequence 8, Appli |
| 42 | 43 | 26.7 | 47 | 1 | US-08-415-751-12 | Sequence 12, Appl |
| 43 | 43 | 26.7 | 116 | 2 | US-08-888-366-12 | Sequence 12, Appl |
| 44 | 43 | 26.7 | 149 | 4 | US-08-679-493A-85 | Sequence 85, Appl |
| 45 | 43 | 26.7 | 231 | 3 | US-08-721-259-2 | Sequence 2, Appli |

ALIGNMENTS

```
RESULT 1
US-08-193-114B-1
; Sequence 1, Application US/08193114B
; Patent No. 5472945
; GENERAL INFORMATION:
;   APPLICANT: Schmaier, Alvin H.
;   APPLICANT: Jiang, Yongping
;   TITLE OF INVENTION: Modulation of Blood
;   TITLE OF INVENTION: Pressure and Inhibition of Platelet Activation
;   TITLE OF INVENTION: with Kininogen Fragment
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Seidel, Gonda, Lavorgna &
;     ADDRESSEE: Monaco, P.C.
;     STREET: 1800 Two Penn Center Plaza
;     CITY: Philadelphia
;     STATE: Pennsylvania
;     COUNTRY: U.S.A.
;     ZIP: 19102
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
;     COMPUTER: IBM PS/2
;     OPERATING SYSTEM: MS-DOS
;     SOFTWARE: WordPerfect 5.1
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/193,114B
;     FILING DATE: 9 February 1994
;     CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: U.S. Application
;     APPLICATION NUMBER: Serial No. 5472945 07/744,545
;     FILING DATE: 13 August 1991
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Monaco, Daniel A.
;     REGISTRATION NUMBER: 30,480
;     REFERENCE/DOCKET NUMBER: 6056-137 CIl
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (215) 568-8383
;     TELEFAX: (215) 568-5549
;     TELEX: No. 5472945e
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 117 amino acids
;     TYPE: peptide
;     TOPOLOGY: linear
US-08-193-114B-1


  Query Match            100.0%;  Score 161;  DB 1;  Length 117;
  Best Local Similarity  100.0%;  Pred. No. 4.9e-15;
```

```
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32
          ||||||||||||||||||||||||||||||||
Db     18 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 49


RESULT   2
PCT-US92-06809-1
; Sequence 1, Application PC/TUS9206809
;  GENERAL INFORMATION:
;    APPLICANT:  Schmaier, Alvin H.
;    APPLICANT:  Jiang, Yongping
;    TITLE OF INVENTION:  Modulation of Blood
;    TITLE OF INVENTION:  Pressure by Altering Bradykinin Levels
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of the
;      ADDRESSEE:  Commonwealth System of Higher Education
;      STREET:  406 University Services
;      STREET:  Building
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/06809
;      FILING DATE:  19910813
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  U.S. Application
;      APPLICATION NUMBER:  Serial No.  744,545
;      FILING DATE:  13 August 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-137
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  117 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
PCT-US92-06809-1


  Query Match             100.0%;  Score 161;  DB 5;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-15;
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32
          ||||||||||||||||||||||||||||||||
Db     18 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 49


RESULT   3
US-08-676-242-15
; Sequence 15, Application US/08676242C
; Patent No. 6143719
; GENERAL INFORMATION:
;  APPLICANT: The Regents of the University of Michigan
;  APPLICANT: Schmaier, Alvin H.
;  APPLICANT: Hasan, Ahmed A.K.
;  TITLE OF INVENTION: Bradykinin Analogs As Selective Thrombin Inhibitors
;  FILE REFERENCE: 8820-2 US
;  CURRENT APPLICATION NUMBER: US/08/676,242C
;  CURRENT FILING DATE: 2000-07-16
;  EARLIER APPLICATION NUMBER: 60/000,096
;  EARLIER FILING DATE: 1995-06-09
;  EARLIER APPLICATION NUMBER: PCT/US96/09940
;  EARLIER FILING DATE: 1996-06-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 26
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Bradykinin
;   OTHER INFORMATION: analog
US-08-676-242-15


  Query Match             59.0%;  Score 95;  DB 4;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 9.1e-07;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     14 NATFYFKIDNVKKARVQVV 32
          |||||||||||||||||||
Db      1 NATFYFKIDNVKKARVQVV 19


RESULT   4
US-08-533-669A-16
; Sequence 16, Application US/08533669A
; Patent No. 5834592
;  GENERAL INFORMATION:
;    APPLICANT:  Corixa Corporation
;    TITLE OF INVENTION:  LEISHMANIA ANTIGENS FOR USE IN THE
;    TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF LEISHMANIASIS
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/533,669A
;      FILING DATE:  22-SEP-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.420
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  701 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-533-669A-16


  Query Match             32.3%;  Score 52;  DB 2;  Length 701;
  Best Local Similarity   38.5%;  Pred. No. 24;
  Matches   10;  Conservative    6;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLTHTITKLNAENNATFYFKIDNVKK 26
```

```
          ||   :|:: || |: :|   |: ||
Db    448 TLKDYVTRMKAEQNSIYYITGDSKKK 473


RESULT   5
US-08-483-926A-11
; Sequence 11, Application US/08483926A
; Patent No. 5821227
;  GENERAL INFORMATION:
;    APPLICANT:  Dennis, James W.
;    TITLE OF INVENTION:  MODULATORS OF CYTOKINES OF THE TGF BETA
;    TITLE OF INVENTION:  SUPERFAMILY AND METHODS FOR ASSAYING FOR SAME
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BERESKIN & PARR
;      STREET:  40 king Street West
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5H 3Y2
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,926A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kurdydyk, Linda M.
;      REGISTRATION NUMBER:  34,971
;      REFERENCE/DOCKET NUMBER:  3153-155
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 364-7311
;      TELEFAX:  (416) 361-1398
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  352 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Rat
US-08-483-926A-11


  Query Match             30.7%;  Score 49.5;  DB 2;  Length 352;
  Best Local Similarity   35.7%;  Pred. No. 25;
  Matches   10;  Conservative    7;  Mismatches   10;  Indels    1;  Gaps    1;

Qy       1 TLTHTITKLNAENNATFYFKIDNVKKAR 28
           |:   :   ||:|| | |||:  : :|:
Db     163 TVKTALAAFNAQNNGT-YFKLVEISRAQ 189


RESULT   6
US-08-737-045-12
; Sequence 12, Application US/08737045A
; Patent No. 5981483
; GENERAL INFORMATION:
;  APPLICANT: Dennis, James W.
;  APPLICANT: Denetriou, Michael
;  APPLICANT: Mount Sinai Hospital Corporation
;  TITLE OF INVENTION: COMPOSITIONS COMPRISING MODULATORS OF CYTOKINES OF THE
;  TITLE OF INVENTION: TGFb SUPERFAMILY AND A METHOD OF TREATMENT WITH SUCH A
;  TITLE OF INVENTION: COMPOSITION (AS AMENDED)
;  FILE REFERENCE: 7933.94USWO
;  CURRENT APPLICATION NUMBER: US/08/737,045A
;  CURRENT FILING DATE: 1997-03-20
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 352
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism:Rat
US-08-737-045-12


  Query Match             30.7%;  Score 49.5;  DB 2;  Length 352;
  Best Local Similarity   35.7%;  Pred. No. 25;
  Matches   10;  Conservative    7;  Mismatches   10;  Indels    1;  Gaps    1;

Qy       1 TLTHTITKLNAENNATFYFKIDNVKKAR 28
           |:   :   ||:|| | |||:  : :|:
Db     163 TVKTALAAFNAQNNGT-YFKLVEISRAQ 189


RESULT   7
US-08-832-535-2
; Sequence 2, Application US/08832535
; Patent No. 5919658
;  GENERAL INFORMATION:
;    APPLICANT:  NI, JIAN
;    APPLICANT:  LI, HAODONG
;    APPLICANT:  YU, GUO-LIANG
;    APPLICANT:  GENTZ, REINER L
;    TITLE OF INVENTION:  HUMAN CYSTATIN F
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HUMAN GENOME SCIENCES, INC.
;      STREET:  9410 KEY WEST AVENUE
;      CITY:  ROCKVILLE
;      STATE:  MD
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/832,535
;      FILING DATE:  03-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KIMBALL, PAUL C.
;      REGISTRATION NUMBER:  34,610
;      REFERENCE/DOCKET NUMBER:  PF265
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (201) 994-1700
;      TELEFAX:  (201) 994-1744
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  145 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-832-535-2


  Query Match             29.8%;  Score 48;  DB 2;  Length 145;
  Best Local Similarity   34.5%;  Pred. No. 15;
  Matches   10;  Conservative    6;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 HTITKLNAENNATFYFKIDNVKKARVQVV 32
           ::: | |   |  | ||   : :| ||:|
Db      55 YSVEKFNNCTNDMFLFKESRITRALVQIV 83


RESULT   8
US-09-019-485-2
; Sequence 2, Application US/09019485
```

```
; Patent No. 6066617
;  GENERAL INFORMATION:
;    APPLICANT:  Li, Haodong
;    APPLICANT:  Yu, Guo-Liang
;    APPLICANT:  Gentz, Reiner
;    APPLICANT:  Ni, Jian
;    TITLE OF INVENTION:  Cystatin F
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/019,485
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Robert H.
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PF265P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  3013098504
;      TELEFAX:  3013098439
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  145 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-019-485-2


  Query Match             29.8%;  Score 48;  DB 3;  Length 145;
  Best Local Similarity   34.5%;  Pred. No. 15;
  Matches   10;  Conservative    6;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 HTITKLNAENNATFYFKIDNVKKARVQVV 32
           ::: | |    |  | ||   : :| ||:|
Db      55 YSVEKFNNCTNDMFLFKESRITRALVQIV 83


RESULT   9
US-09-019-485-3
; Sequence 3, Application US/09019485
; Patent No. 6066617
;  GENERAL INFORMATION:
;    APPLICANT:  Li, Haodong
;    APPLICANT:  Yu, Guo-Liang
;    APPLICANT:  Gentz, Reiner
;    APPLICANT:  Ni, Jian
;    TITLE OF INVENTION:  Cystatin F
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/019,485
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Robert H.
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PF265P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  3013098504
;      TELEFAX:  3013098439
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  145 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-019-485-3


  Query Match             29.8%;  Score 48;  DB 3;  Length 145;
  Best Local Similarity   34.5%;  Pred. No. 15;
  Matches   10;  Conservative    6;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 HTITKLNAENNATFYFKIDNVKKARVQVV 32
           ::: | |    |  | ||   : :| ||:|
Db      55 YSVEKFNNCTNDMFLFKESRITRALVQIV 83


RESULT  10
US-09-431-480-9
; Sequence 9, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
;  APPLICANT: Holloway, James L.
;  APPLICANT: Feldhaus, Andrew
;  TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
;  FILE REFERENCE: 98-72
;  CURRENT APPLICATION NUMBER: US/09/431,480
;  CURRENT FILING DATE: 1999-11-01
;  EARLIER APPLICATION NUMBER: 60/109,217
;  EARLIER FILING DATE: 1998-11-20
;  EARLIER APPLICATION NUMBER: 60/156,382
;  EARLIER FILING DATE: 1999-09-28
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9
;   LENGTH: 145
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-431-480-9


  Query Match             29.8%;  Score 48;  DB 4;  Length 145;
  Best Local Similarity   34.5%;  Pred. No. 15;
  Matches   10;  Conservative    6;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 HTITKLNAENNATFYFKIDNVKKARVQVV 32
           ::: | |    |  | ||   : :| ||:|
Db      55 YSVEKFNNCTNDMFLFKESRITRALVQIV 83


RESULT  11
US-09-617-302-9
; Sequence 9, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
;  APPLICANT:  Holloway, James L.
;  APPLICANT:  Feldhaus, Andrew
;  TITLE OF INVENTION:  TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
;  FILE REFERENCE:  98-72 C1
;  CURRENT APPLICATION NUMBER: US/09/617,302
;  CURRENT FILING DATE:  2000-07-17
```

```
;  PRIOR APPLICATION NUMBER:  09/431,480
;  PRIOR FILING DATE:  1999-11-01
;  PRIOR APPLICATION NUMBER:  60/109,217
;  PRIOR FILING DATE:  1998-11-20
;  PRIOR APPLICATION NUMBER:  60/156,382
;  PRIOR FILING DATE:  1999-09-28
;  NUMBER OF SEQ ID NOS:  22
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 9
;   LENGTH: 145
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-617-302-9


  Query Match             29.8%;  Score 48;  DB 4;  Length 145;
  Best Local Similarity   34.5%;  Pred. No. 15;
  Matches   10;  Conservative    6;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 HTITKLNAENNATFYFKIDNVKKARVQVV 32
           ::: | |   |  | ||   : :| ||:|
Db      55 YSVEKFNNCTNDMFLFKESRITRALVQIV 83


RESULT  12
US-08-791-522-1
; Sequence 1, Application US/08791522
; Patent No. 5935817
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Goli, Surya K.
;    TITLE OF INVENTION:  NOVEL HUMAN CYSTATIN-LIKE
;    TITLE OF INVENTION:  PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/791,522
;      FILING DATE:  Filed Herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0193 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  178 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  30443
US-08-791-522-1


  Query Match             29.8%;  Score 48;  DB 2;  Length 178;
  Best Local Similarity   34.5%;  Pred. No. 19;
  Matches   10;  Conservative    6;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 HTITKLNAENNATFYFKIDNVKKARVQVV 32
           ::: | |   |  | ||   : :| ||:|
Db      77 YSVEKFNNCTNDMFLFKESRITRALVQIV 105


RESULT  13
US-09-314-777-1
; Sequence 1, Application US/09314777
; Patent No. 6110686
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Goli, Surya K.
;    TITLE OF INVENTION:  NOVEL HUMAN CYSTATIN-LIKE
;    TITLE OF INVENTION:  PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/314,777
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/791,522
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0193 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  178 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  30443
US-09-314-777-1


  Query Match             29.8%;  Score 48;  DB 3;  Length 178;
  Best Local Similarity   34.5%;  Pred. No. 19;
  Matches   10;  Conservative    6;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 HTITKLNAENNATFYFKIDNVKKARVQVV 32
           ::: | |   |  | ||   : :| ||:|
Db      77 YSVEKFNNCTNDMFLFKESRITRALVQIV 105


RESULT  14
US-08-415-751-20
; Sequence 20, Application US/08415751
; Patent No. 5643772
;  GENERAL INFORMATION:
;    APPLICANT:  PETERSEN, CAROLYN
;    APPLICANT:  LEECH, JAMES
;    APPLICANT:  NELSON, RICHARD, C.
;    APPLICANT:  GUT, JIRI
;    TITLE OF INVENTION:  POLYPEPTIDES BINDING ANTI-
```

```
;    TITLE OF INVENTION:  CRYPTOSPORIDIUM ANTIBODIES, DNA
;    TITLE OF INVENTION:  AND RNA ENCODING THEM, HYBRID
;    TITLE OF INVENTION:  VECTOR AND TRANSFORMED HOST AND
;    TITLE OF INVENTION:  METHODS FOR IMMUNOTHERAPY AND
;    TITLE OF INVENTION:  DIAGNOSIS AND KIT
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILLIPS, MOORE, LEMPIO & FINLEY
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette - 3.5 inch, 1.44 Kb storage
;      COMPUTER:  PC
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  Wordperfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/415,751
;      FILING DATE:  03-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/071,880
;      FILING DATE:  June 1, 1993
;      APPLICATION NUMBER:  07/891,301
;      FILING DATE:  May 29, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hana Dolezalova
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  480.19-2 (HHD)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-1677
;      TELEFAX:  (415) 324-1678
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  350 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Cryptosporidium parvum
;    FEATURE:
;      NAME/KEY:  Positions coded by nonsense codons are
;      NAME/KEY:  identified as Xaa.
US-08-415-751-20


  Query Match             29.8%;  Score 48;  DB 1;  Length 350;
  Best Local Similarity   39.1%;  Pred. No. 40;
  Matches    9;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       4 HTITKLNAENNATFYFKIDNVKK 26
           ||  | |   | | || : ::|:
Db     244 HTNVKFNGXKNNTTYFPLXSLKE 266


RESULT  15
US-08-751-359-11
; Sequence 11, Application US/08751359
; Patent No. 6143559
;  GENERAL INFORMATION:
;    APPLICANT:  Michael, Nancy M
;    APPLICANT:  Accavitti, Marianne
;    APPLICANT:  Thompson, Craig B
;    TITLE OF INVENTION:  METHODS FOR THE PRODUCTION OF CHICKEN
;    TITLE OF INVENTION:  MONOCLONAL ANTIBODIES
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/751,359
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  510
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  ARSB:504
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  126 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-751-359-11


  Query Match             28.6%;  Score 46;  DB 4;  Length 126;
  Best Local Similarity   50.0%;  Pred. No. 24;
  Matches    9;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAENNATFY 18
           | | ||| : |::|| :|
Db      86 TATLTITGVRADDNAVYY 103


Search completed: July  1, 2002, 16:18:02
Job time: 46 sec
```

GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:19:44 ; Search time 95.97 Seconds
(without alignments)
37.036 Million cell updates/sec

Title: US-09-461-061A-4
Perfect score: 161
Sequence: 1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 161 | 100.0 | 32 | 21 | AAY95408 | Anti-angiogenic D3 |
| 2 | 161 | 100.0 | 117 | 14 | AAR33350 | Domaine 3, bradyki |
| 3 | 161 | 100.0 | 122 | 21 | AAB37447 | Human kininogen D3 |
| 4 | 161 | 100.0 | 123 | 21 | AAY95426 | Human high mol.wt. |
| 5 | 161 | 100.0 | 644 | 22 | ABG21101 | Novel human diagno |
| 6 | 157 | 97.5 | 369 | 22 | ABG21099 | Novel human diagno |
| 7 | 153 | 95.0 | 248 | 22 | ABG21102 | Novel human diagno |
| 8 | 143 | 88.8 | 435 | 22 | ABG21105 | Novel human diagno |
| 9 | 116 | 72.0 | 434 | 5 | AAP40633 | Bradykinin protein |
| 10 | 112 | 69.6 | 436 | 5 | AAP40257 | Bradykinin protein |
| 11 | 95 | 59.0 | 26 | 18 | AAW54336 | Bradykinin analogo |
| 12 | 84 | 52.2 | 16 | 21 | AAY95410 | Anti-angiogenic D3 |
| 13 | 83 | 51.6 | 16 | 21 | AAY95409 | Anti-angiogenic D3 |
| 14 | 74 | 46.0 | 16 | 21 | AAB08553 | Peptide identified |
| 15 | 61 | 37.9 | 167 | 20 | AAW98907 | Mouse IMC carcinom |
| 16 | 60 | 37.3 | 12 | 21 | AAY95406 | Anti-angiogenic pe |
| 17 | 56 | 34.8 | 12 | 21 | AAY95407 | Anti-angiogenic pe |
| 18 | 52 | 32.3 | 316 | 22 | ABB62195 | Drosophila melanog |
| 19 | 51 | 31.7 | 121 | 21 | AAY81200 | Human mutant cysta |
| 20 | 51 | 31.7 | 128 | 21 | AAY81189 | Human mutant cysta |
| 21 | 50 | 31.1 | 118 | 21 | AAY81218 | Bovine mutant cyst |
| 22 | 50 | 31.1 | 121 | 21 | AAY81198 | Human mutant cysta |
| 23 | 50 | 31.1 | 128 | 21 | AAY81187 | Human mutant cysta |
| 24 | 50 | 31.1 | 209 | 21 | AAG58043 | Arabidopsis thalia |
| 25 | 50 | 31.1 | 218 | 21 | AAG58042 | Arabidopsis thalia |
| 26 | 49 | 30.4 | 9 | 21 | AAB37455 | Human kininogen D3 |
| 27 | 48 | 29.8 | 75 | 19 | AAW62818 | Amino acid sequenc |
| 28 | 48 | 29.8 | 75 | 19 | AAW62802 | Amino acid sequenc |
| 29 | 48 | 29.8 | 126 | 21 | AAB37445 | Human cystatin F. |
| 30 | 48 | 29.8 | 145 | 18 | AAW32323 | Mature human cysta |
| 31 | 48 | 29.8 | 145 | 18 | AAW31902 | Human cystatin F. |
| 32 | 48 | 29.8 | 145 | 20 | AAY25708 | Human cystatin F p |
| 33 | 48 | 29.8 | 145 | 22 | AAE04439 | Human cystatin F h |
| 34 | 48 | 29.8 | 145 | 22 | AAE02410 | Human cystatin F p |
| 35 | 48 | 29.8 | 167 | 20 | AAY02287 | Secreted protein c |
| 36 | 48 | 29.8 | 167 | 20 | AAW98910 | Mouse IMC carcinom |
| 37 | 48 | 29.8 | 178 | 19 | AAW69734 | Human cystatin-lik |
| 38 | 48 | 29.8 | 271 | 22 | AAG82864 | S. epidermidis ope |
| 39 | 48 | 29.8 | 271 | 22 | AAG83107 | S. epidermidis ope |
| 40 | 48 | 29.8 | 353 | 20 | AAY19859 | B. burgdorferi ant |
| 41 | 48 | 29.8 | 373 | 20 | AAY19858 | B. burgdorferi ant |
| 42 | 48 | 29.8 | 611 | 22 | AAU33236 | Novel human secret |
| 43 | 48 | 29.8 | 695 | 22 | AAU33237 | Novel human secret |
| 44 | 47.5 | 29.5 | 144 | 22 | ABB69396 | Drosophila melanog |
| 45 | 47.5 | 29.5 | 452 | 22 | ABB50268 | Mitochondrial EF-T |

ALIGNMENTS

RESULT 1
AAY95408
ID AAY95408 standard; Peptide; 32 AA.
XX
AC AAY95408;
XX
DT 25-SEP-2000 (first entry)
XX
DE Anti-angiogenic D3 peptide.
XX
KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW therapy; human; D3 peptide.
XX
OS Homo sapiens.
XX
PN WO200035407-A2.
XX
PD 22-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28465.
XX
PR 16-DEC-1998; 98US-0112427.
XX
PA (UTEM ) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI McCrae RK;
XX
DR WPI; 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,

PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT 3 analog -
XX
PS Claim 4; Page 26; 44pp; English.
XX
CC The present sequence is that of a D3 peptide derived from human
CC high mol.wt. kininogen (HK) domain 3 (see AAY95426). The D3 peptide
CC inhibits endothelial cell proliferation and thus possesses
CC anti-angiogenic activity. It is an example of D3 peptides of the
CC invention (see AAY95405-26) that are analogues of certain sites in
CC the HK domain 3, in this case amino acid residues Asn275-Lys282.
CC The peptides inhibit endothelial cell proliferation and may also
CC induce endothelial cell apoptosis. Compositions including the
CC peptides are used in claimed methods for inhibiting angiogenesis,
CC inhibiting endothelial cell proliferation, and inducing endothelial
CC cell apoptosis. Cancer, rheumatoid arthritis, and ocular disorders
CC characterized by undesired vascularization of the retina are treated.
XX
SQ Sequence 32 AA;

Query Match 100.0%; Score 161; DB 21; Length 32;
Best Local Similarity 100.0%; Pred. No. 3.6e-15;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32
||||||||||||||||||||||||||||||||
Db 1 tlthtitklnaennatfyfkidnvkkarvqvv 32

RESULT 2
AAR33350
ID AAR33350 standard; protein; 117 AA.
XX
AC AAR33350;
XX
DT 01-JUL-1993 (first entry)
XX
DE Domaine 3, bradykinin release activating peptide.
XX
KW Domain 3; human; kininogen; heavy chain; low molecular weight; plasma;
KW trypsin; platelet; activation; granule contents; hemostasis; thrombin;
KW tissue plasminogen activator;thrombosis; inflammatory response;
KW endothelial cell; von Willebrand factor;
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..18
FT /note= "Leader peptide"
FT Protein 19..117
FT /note= "Mature protein"
XX
PN WO9303748-A.
XX
PD 04-MAR-1993.
XX
PF 13-AUG-1992; 92WO-US06809.
XX
PR 13-AUG-1991; 91US-0744545.
XX
PA (UTEM ) UNIV TEMPLE.
XX
PI Jiang Y, Schmaier AB;
XX
DR WPI; 1993-093714/11.
XX
PT Use of trypsin-cleavage fragment of human kininogen - for
PT increasing vascular bradykinin release, for lowering blood
PT pressure and treating hypertension
XX
PS Disclosure; Fig 1; 46pp; English.
XX
CC The sequence given represents domain 3, amino acids 246-362, of
CC the human kininogen heavy chain. Domain 3 was isolated from low
CC molecular weight kininogen, derived from human plasma, by cleavage
CC with trypsin. Domain 3 peptide inhibits platelet activation causing
CC a marked decrease in the platelets ability to aggregate and secrete
CC their granule contents. The granule contents comprise proteins which
CC participate in hemostasis, thrombosis and the inflammatory response.
CC Domain 3 also inhibits endothelial cell activation shown by a decrease
CC in secretion of endothelial cell contents such as tissue plasminogen
CC activator and von Willebrand factor. Domain 3 functions to inhibit
CC cell activation by blocking thrombin binding to its target cells, the
CC peptide is a selective inhibitor of thrombin-induced platlet
CC activation.
XX
SQ Sequence 117 AA;

Query Match 100.0%; Score 161; DB 14; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.6e-14;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32
||||||||||||||||||||||||||||||||
Db 18 tlthtitklnaennatfyfkidnvkkarvqvv 49

not 0-12

RESULT 3
AAB37447
ID AAB37447 standard; protein; 122 AA.
XX
AC AAB37447;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human kininogen D3.
XX
KW Enzyme; legumain; endopeptidase; cystatin; human; kininogen.
XX
OS Homo sapiens.
XX
PN WO200064945-A1.
XX
PD 02-NOV-2000.
XX
PF 20-APR-2000; 2000WO-GB01571.
XX
PR 22-APR-1999; 99GB-0009133.
XX
PA (BABR-) BABRAHAM INST.
XX
PI Abrahamson M, Barrett AJ;
XX
DR WPI; 2000-687316/67.
XX
PT Inhibition of mammalian legumain or legumain-related endopeptidase by
PT cystatin involves interaction with second papain-non-reactive site of
PT cystatin -
XX
PS Disclosure; Fig 4; 45pp; English.
XX
CC The present invention relates to inhibition of the enzymatic activity of
CC legumain or a legumain-related endopeptidase by cystatin. The inhibition
CC involves an interaction between legumain and a papain-non-reactive site
CC of cystatin. Legumain (EC 3.4.22.34) is a cysteine endopeptidase, and
CC performs a protein-processing function. The present sequence is human
CC kininogen D3, which was used in the present invention. Kininogen is a
CC type 3 cystatin.
XX
SQ Sequence 122 AA;

Query Match 100.0%; Score 161; DB 21; Length 122;
Best Local Similarity 100.0%; Pred. No. 1.7e-14;

```
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           ||||||||||||||||||||||||||||||||
Db      23 tlthtitklnaennatfyfkidnvkkarvqvv 54


RESULT   4
AAY95426
ID   AAY95426 standard; Peptide; 123 AA.
XX
AC   AAY95426;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Human high mol.wt. kininogen domain 3.
XX
KW   Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW   endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW   rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW   therapy; human; D3 peptide.
XX
OS   Homo sapiens.
XX
PN   WO200035407-A2.
XX
PD   22-JUN-2000.
XX
PF   02-DEC-1999;   99WO-US28465.
XX
PR   16-DEC-1998;   98US-0112427.
XX
PA   (UTEM ) UNIV TEMPLE.
PA   (MCCR/) MCCRAE R K.
XX
PI   McCrae RK;
XX
DR   WPI; 2000-442247/38.
XX
PT   Composition for inhibiting angiogenesis and endothelial cell
PT   proliferation, inducing endothelial cell apoptosis and treating cancer,
PT   rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT   3 analog  -
XX
PS   Disclosure; Page 4; 44pp; English.
XX
CC   The present sequence is that of domain 3 of human high mol.wt.
CC   kininogen (HK).  The invention provides peptides (see AAY95405-24)
CC   that are analogues of certain sites in the HK domain 3,
CC   specifically Asn275-Lys282, Cys246-Cys249, Leu331-Tyr338 and
CC   Tyr299-Ser314.  The peptides, in which native Cys residues may be
CC   replaced by Ala residues, inhibit endothelial cell proliferation
CC   and may also induce endothelial cell apoptosis.  Compositions
CC   including the peptides are used in claimed methods for inhibiting
CC   angiogenesis, inhibiting endothelial cell proliferation, and
CC   inducing endothelial cell apoptosis.  Cancer, rheumatoid arthritis,
CC   and ocular disorders characterized by undesired vascularization of
CC   the retina are treated.
XX
SQ   Sequence   123 AA;


  Query Match           100.0%;  Score 161;  DB 21;  Length 123;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-14;
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           ||||||||||||||||||||||||||||||||
Db      29 tlthtitklnaennatfyfkidnvkkarvqvv 60


RESULT   5
ABG21101
ID   ABG21101 standard; Protein; 644 AA.
XX
AC   ABG21101;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #21092.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS85288.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 51460; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   644 AA;


  Query Match           100.0%;  Score 161;  DB 22;  Length 644;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-13;
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           ||||||||||||||||||||||||||||||||
Db     281 tlthtitklnaennatfyfkidnvkkarvqvv 312


RESULT   6
ABG21099
ID   ABG21099 standard; Protein; 369 AA.
```

```
XX
AC   ABG21099;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #21090.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS85286.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 51458; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   369 AA;


  Query Match             97.5%;  Score 157;  DB 22;  Length 369;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-13;
  Matches   31;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAENNATFYFKIDNVKKARVQV 31
           |||||||||||||||||||||||||||||||
Db     321 tlthtitklnaennatfyfkidnvkkarvqv 351


RESULT  7
ABG21102
ID   ABG21102 standard; Protein; 248 AA.
XX
```

```
AC   ABG21102;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #21093.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS85289.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 51461; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   248 AA;


  Query Match             95.0%;  Score 153;  DB 22;  Length 248;
  Best Local Similarity   100.0%;  Pred. No. 4.7e-13;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAENNATFYFKIDNVKKARVQ 30
           ||||||||||||||||||||||||||||||
Db      78 tlthtitklnaennatfyfkidnvkkarvq 107


RESULT  8
ABG21105
ID   ABG21105 standard; Protein; 435 AA.
XX
AC   ABG21105;
```

```
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #21096.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS85292.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 51464; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   435 AA;


  Query Match             88.8%;  Score 143;  DB 22;  Length 435;
  Best Local Similarity   90.6%;  Pred. No. 2.1e-11;
  Matches   29;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           ||||||||||||||  ||||||||||||:||
Db     285 tlthtitklnaennaisyfkidnvkkarvrvv 316


RESULT   9
AAP40633
ID   AAP40633 standard; Protein; 434 AA.
XX
AC   AAP40633;
XX
DT   30-JUL-1992  (first entry)
XX
DE   Bradykinin protein precursor: type II (pKG146, pKG254).
XX
KW   Blood pressure; kininogen; probe..
XX
FH   Key             Location/Qualifiers
FT   Peptide         378..386
FT                   /label= bradykinin
FT   Peptide         391..395
FT                   /note= "probe (AAN40241)-encoded sequence"
XX
PN   JP59125896-A.
XX
PD   20-JUL-1984.
XX
PF   07-JAN-1983;   83JP-0000984.
XX
PR   07-JAN-1983;   83JP-0000984.
XX
PA   (MITU ) MITSUBISHI CHEM IND KK.
XX
DR   WPI; 1984-216122/35.
DR   N-PSDB; AAN40314.
XX
PT   c-Dna fragment of protein precursor - used to code bradykinin
XX
PS   Disclosure; Fig 2; 6 pp; Japanese.
XX
CC   Bradykinin is a peptide consisting of nine amino acids.  It has the
CC   biological effect of decreasing blood pressure.  Although kininogen
CC   is known as a protein-precursor of bradykinin, its structure is unknown
CC   because of the difficulty in collecting large enough samples of
CC   kininogen for structural investigation.
XX
SQ   Sequence   434 AA;


  Query Match             72.0%;  Score 116;  DB 5;  Length 434;
  Best Local Similarity   74.2%;  Pred. No. 1e-07;
  Matches   23;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           | |:| ||||||:: ||||||| |||| ||||
Db     279 lnhsiaklnaehdgtfyfkidtvkkatvqvv 309


RESULT  10
AAP40257
ID   AAP40257 standard; Protein; 436 AA.
XX
AC   AAP40257;
XX
DT   30-JUL-1992  (first entry)
XX
DE   Bradykinin protein precursor: type I (pKG13, pKG59).
XX
KW   Blood pressure; kininogen; probe.
XX
FH   Key             Location/Qualifiers
FT   Peptide         380..388
FT                   /label= bradykinin
FT   Peptide         393..397
FT                   /note= "probe (AAN40241)-encoded sequence"
XX
PN   JP59125896-A.
XX
PD   20-JUL-1984.
XX
PF   07-JAN-1983;   83JP-0000984.
XX
PR   07-JAN-1983;   83JP-0000984.
XX
```

```
PA   (MITU ) MITSUBISHI CHEM IND KK.
XX
DR   WPI; 1984-216122/35.
DR   N-PSDB; AAN40242.
XX
PT   c-Dna fragment of protein precursor - used to code bradykinin
XX
PS   Disclosure; Fig 2; 6 pp; Japanese.
XX
CC   Bradykinin is a peptide consisting of nine amino acids.  It has the
CC   biological effect of decreasing blood pressure.  Although kininogen
CC   is known as a protein-precursor of bradykinin, its structure is unknown
CC   because of the difficulty in collecting large enough samples of
CC   kininogen for structural investigation.
XX
SQ   Sequence   436 AA;


  Query Match             69.6%;  Score 112;  DB 5;  Length 436;
  Best Local Similarity   71.0%;  Pred. No. 3.6e-07;
  Matches   22;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           |:|:| |||||::  |||||| |||| ||||
Db     281 lshsiaklnaehdgafyfkidtvkkatvqvv 311


RESULT  11
AAW54336
ID   AAW54336 standard; peptide; 26 AA.
XX
AC   AAW54336;
XX
DT   30-JUL-1998  (first entry)
XX
DE   Bradykinin analogous peptide 19.
XX
KW   Inhibition; thrombin-induced platelet; prevention; platelet aggregation;
KW   ADP-induced activation.
XX
OS   Synthetic.
XX
PN   WO9641640-A1.
XX
PD   27-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US09940.
XX
PR   09-JUN-1995;   95US-0000096.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Hasan AAK,  Schmaier AH;
XX
DR   WPI; 1997-065304/06.
XX
PT   Inhibition of platelet activation and aggregation - by admin. of new
PT   or known bradykinin analogues
XX
PS   Disclosure; Page 44; 73pp; English.
XX
CC   Administration of a peptide or multimer related to bradykinin or other
CC   disclosed peptides and multimers can be used for the inhibition of
CC   thrombin-induced platelets or other cells.  They can also be used for
CC   preventing platelet aggregation, or inhibiting ADP-induced activation.
CC   This is useful to prevent arterial occlusions arising from coronary
CC   thrombosis and stroke.
XX
SQ   Sequence   26 AA;


  Query Match             59.0%;  Score 95;  DB 18;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-06;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      14 NATFYFKIDNVKKARVQVV 32
           ||||||||||||||||||
Db       1 natfyfkidnvkkarvqvv 19


RESULT  12
AAY95410
ID   AAY95410 standard; Peptide; 16 AA.
XX
AC   AAY95410;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Anti-angiogenic D3 peptide.
XX
KW   Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW   endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW   rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW   therapy; human; D3 peptide.
XX
OS   Homo sapiens.
XX
PN   WO200035407-A2.
XX
PD   22-JUN-2000.
XX
PF   02-DEC-1999;   99WO-US28465.
XX
PR   16-DEC-1998;   98US-0112427.
XX
PA   (UTEM ) UNIV TEMPLE.
PA   (MCCR/) MCCRAE R K.
XX
PI   McCrae RK;
XX
DR   WPI; 2000-442247/38.
XX
PT   Composition for inhibiting angiogenesis and endothelial cell
PT   proliferation, inducing endothelial cell apoptosis and treating cancer,
PT   rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT   3 analog  -
XX
PS   Claim 7; Page 26; 44pp; English.
XX
CC   The present sequence is that of a D3 peptide derived from human
CC   high mol.wt. kininogen (HK) domain 3 (see AAY95426).  The D3 peptide
CC   inhibits endothelial cell proliferation and thus possesses
CC   anti-angiogenic activity.  It is an example of D3 peptides of the
CC   invention (see AAY95405-26) that are analogues of certain sites in
CC   the HK domain 3, in this case amino acid residues Asn275-Lys282.
CC   The peptides inhibit endothelial cell proliferation and may also
CC   induce endothelial cell apoptosis.  Compositions including the
CC   peptides are used in claimed methods for inhibiting angiogenesis,
CC   inhibiting endothelial cell proliferation, and inducing endothelial
CC   cell apoptosis.  Cancer, rheumatoid arthritis, and ocular disorders
CC   characterized by undesired vascularization of the retina are treated.
CC   The IC50 value for the present peptide was less than 0.8 uM for
CC   inhibition of fibroblast growth factor-induced HUVEC cell
CC   proliferation.
XX
SQ   Sequence   16 AA;


  Query Match             52.2%;  Score 84;  DB 21;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 NNATFYFKIDNVKKAR 28
           ||||||||||||||||
Db       1 nnatfyfkidnvkkar 16
```

```
RESULT 13
AAY95409
ID   AAY95409 standard; Peptide; 16 AA.
XX
AC   AAY95409;
XX
DT   25-SEP-2000 (first entry)
XX
DE   Anti-angiogenic D3 peptide.
XX
KW   Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW   endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW   rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW   therapy; human; D3 peptide.
XX
OS   Homo sapiens.
XX
PN   WO200035407-A2.
XX
PD   22-JUN-2000.
XX
PF   02-DEC-1999;   99WO-US28465.
XX
PR   16-DEC-1998;   98US-0112427.
XX
PA   (UTEM ) UNIV TEMPLE.
PA   (MCCR/) MCCRAE R K.
XX
PI   McCrae RK;
XX
DR   WPI; 2000-442247/38.
XX
PT   Composition for inhibiting angiogenesis and endothelial cell
PT   proliferation, inducing endothelial cell apoptosis and treating cancer,
PT   rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT   3 analog -
XX
PS   Claim 6; Page 26; 44pp; English.
XX
CC   The present sequence is that of a D3 peptide derived from human
CC   high mol.wt. kininogen (HK) domain 3 (see AAY95426). The D3 peptide
CC   inhibits endothelial cell proliferation and thus possesses
CC   anti-angiogenic activity. It is an example of D3 peptides of the
CC   invention (see AAY95405-26) that are analogues of certain sites in
CC   the HK domain 3, in this case amino acid residues Asn275-Lys282.
CC   The peptides inhibit endothelial cell proliferation and may also
CC   induce endothelial cell apoptosis. Compositions including the
CC   peptides are used in claimed methods for inhibiting angiogenesis,
CC   inhibiting endothelial cell proliferation, and inducing endothelial
CC   cell apoptosis. Cancer, rheumatoid arthritis, and ocular disorders
CC   characterized by undesired vascularization of the retina are treated.
CC   The IC50 value for the present peptide was less than 0.8 uM for
CC   inhibition of fibroblast growth factor-induced HUVEC cell
CC   proliferation.
XX
SQ   Sequence   16 AA;


  Query Match             51.6%;  Score 83;  DB 21;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 TITKLNAENNATFYFK 20
           ||||||||||||||||
Db       1 titklnaennatfyfk 16


RESULT 14
AAB08553
ID   AAB08553 standard; Peptide; 16 AA.
XX
AC   AAB08553;
XX
DT   20-DEC-2000 (first entry)
XX
DE   Peptide identified from an origin of prepro-bradykinine.
XX
KW   Precursor peptide; polypeptide hormone; peptide identification.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "hydrogen attached"
FT   Modified-site   16
FT                   /note= "amidated residue"
XX
PN   WO200050636-A1.
XX
PD   31-AUG-2000.
XX
PF   24-FEB-2000; 2000WO-FR00460.
XX
PR   25-FEB-1999;   99US-0257525.
XX
PA   (SCRC ) SCRAS SOC CONSEILS RECH & APPL SCI.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Camara Ferrer YJA,  Thurieau C,  Martinez J,  Berge G,  Goze C;
XX
DR   WPI; 2000-572101/53.
XX
PT   Identifying peptide with selected function, useful particularly for
PT   C-amidated hormones, by screening database for combination of nucleic
PT   acid and amino acid sequences -
XX
PS   Claim 16; Page 20; 40pp; French.
XX
CC   The specification describes a method for identifying a peptide having
CC   a particular function. The method comprises preparing a database of
CC   polynucleotides and polypeptides of unknown functions, screening the
CC   database for a combination of nucleotides or amino acids indicative of
CC   the peptide with a particular function, and identifying polynucleotides
CC   and proteins which contain the peptide. The method is used to identify
CC   precursor peptides with an amidated C-terminus, especially polypeptide
CC   hormones, for studying physiologically active substances. The present
CC   sequence represents a peptide which was identified using the method of
CC   the invention.
XX
SQ   Sequence   16 AA;


  Query Match             46.0%;  Score 74;  DB 21;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0012;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      18 YFKIDNVKKARVQVV 32
           |||||||||||||||
Db       1 yfkidnvkkarvqvv 15


RESULT 15
AAW98907
ID   AAW98907 standard; Protein; 167 AA.
XX
AC   AAW98907;
XX
DT   05-MAY-1999 (first entry)
XX
DE   Mouse IMC carcinoma cell IMC-HA1 clone 23-1#2 protein.
XX
KW   Mouse; carcinoma cell; IMC-HA1; cancer; metastasis; CMAP; inhibitor;
KW   cancer metastasis associated protein.
XX
OS   Mus musculus.
```

XX
PN WO9845431-A1.
XX
PD 15-OCT-1998.
XX
PF 07-APR-1998; 98WO-JP01592.
XX
PR 08-APR-1997; 97JP-0105333.
XX
PA (BANY ) BANYU PHARM CO LTD.
XX
PI Arakawa H, Morita M, Ohta M;
XX
DR WPI; 1999-080732/07.
DR N-PSDB; AAX18513.
XX
PT Protein associated with cancer metastasis and gene encoding it -
PT useful for screening for potential inhibitors of cancer metastasis
XX
PS Claim 1; Page 39-40; 74pp; Japanese.
XX
CC The present invention provides gene sequences associated with cancer
CC metastasis which are isolated from mouse IMC carcinoma cells by
CC detection of their higher expression in IMC-HM cell lines than in
CC IMC-LM cell lines using differential display of the mRNA in these cells.
CC The gene sequences can be used for the screening of potential inhibitors
CC of cancer metastasis by either: bringing into contact with the cancer
CC metastasis associated protein (CMAP) and determining the degree of
CC binding; or creating a transformant cell line which expresses CMAP and
CC measuring the degree of expression of CMAP using an antibody recognising
CC the protein, either in the presence or absence of the potential
CC inhibitor. IMC-HM cells transformed with antisense CMAP DNA show a
CC lowered ability to metastasise. The present sequence represents a
CC specifically claimed protein sequence from the present invention.
XX
SQ Sequence 167 AA;

Query Match 37.9%; Score 61; DB 20; Length 167;
Best Local Similarity 48.3%; Pred. No. 1.1;
Matches 14; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

```
Qy       4 HTITKLNAENNATFYFKIDNVKKARVQVV 32
           |:: | |   |  | ||  :| || ||||
Db      77 hsvekfnnctndiflfkeshvskalvqvv 105
```

Search completed: July 1, 2002, 16:19:45
Job time: 149 sec

GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:18:00 ; Search time 35.15 Seconds
(without alignments)
8.339 Million cell updates/sec

Title: US-09-461-061A-3
Perfect score: 56
Sequence: 1 IDNVKKARVQVV 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 56 | 100.0 | 26 | 4 | US-08-676-242-15 | Sequence 15, Appl |
| 2 | 56 | 100.0 | 117 | 1 | US-08-193-114B-1 | Sequence 1, Appli |
| 3 | 56 | 100.0 | 117 | 5 | PCT-US92-06809-1 | Sequence 1, Appli |
| 4 | 36 | 64.3 | 322 | 4 | US-09-359-161-7 | Sequence 7, Appli |
| 5 | 34 | 60.7 | 679 | 1 | US-08-441-139-5 | Sequence 5, Appli |
| 6 | 33 | 58.9 | 607 | 4 | US-09-537-682-1 | Sequence 1, Appli |
| 7 | 32 | 57.1 | 137 | 4 | US-09-188-930-174 | Sequence 174, App |
| 8 | 32 | 57.1 | 360 | 3 | US-08-622-277A-14 | Sequence 14, Appl |
| 9 | 32 | 57.1 | 360 | 4 | US-09-025-580-2 | Sequence 2, Appli |
| 10 | 32 | 57.1 | 380 | 3 | US-08-459-953A-9 | Sequence 9, Appli |
| 11 | 32 | 57.1 | 392 | 3 | US-08-979-917A-2 | Sequence 2, Appli |
| 12 | 32 | 57.1 | 393 | 4 | US-09-252-292C-23 | Sequence 23, Appl |
| 13 | 32 | 57.1 | 423 | 2 | US-08-846-762-74 | Sequence 74, Appl |
| 14 | 32 | 57.1 | 890 | 4 | US-09-342-648-10 | Sequence 10, Appl |
| 15 | 32 | 57.1 | 3898 | 4 | US-08-750-717-2 | Sequence 2, Appli |
| 16 | 31 | 55.4 | 172 | 1 | US-08-487-001A-51 | Sequence 51, Appl |
| 17 | 31 | 55.4 | 172 | 2 | US-08-630-822A-51 | Sequence 51, Appl |
| 18 | 31 | 55.4 | 172 | 2 | US-09-005-069-51 | Sequence 51, Appl |
| 19 | 31 | 55.4 | 254 | 2 | US-08-685-992-23 | Sequence 23, Appl |
| 20 | 31 | 55.4 | 254 | 2 | US-09-144-925-23 | Sequence 23, Appl |
| 21 | 31 | 55.4 | 272 | 2 | US-08-775-009-31 | Sequence 31, Appl |
| 22 | 31 | 55.4 | 365 | 4 | US-09-004-838-133 | Sequence 133, App |
| 23 | 31 | 55.4 | 381 | 2 | US-08-687-355A-4 | Sequence 4, Appli |
| 24 | 31 | 55.4 | 381 | 2 | US-08-687-355A-6 | Sequence 6, Appli |
| 25 | 31 | 55.4 | 398 | 2 | US-08-630-822A-56 | Sequence 56, Appl |
| 26 | 31 | 55.4 | 398 | 2 | US-09-005-069-56 | Sequence 56, Appl |
| 27 | 31 | 55.4 | 456 | 6 | 5432081-7 | Patent No. 5432081 |
| 28 | 31 | 55.4 | 457 | 2 | US-08-882-704A-6 | Sequence 6, Appli |
| 29 | 31 | 55.4 | 457 | 6 | 5268463-7 | Patent No. 5268463 |
| 30 | 31 | 55.4 | 715 | 4 | US-08-669-286-5 | Sequence 5, Appli |
| 31 | 31 | 55.4 | 715 | 4 | US-09-469-253-5 | Sequence 5, Appli |
| 32 | 31 | 55.4 | 715 | 4 | US-09-642-146-5 | Sequence 5, Appli |
| 33 | 31 | 55.4 | 866 | 3 | US-09-040-843-2 | Sequence 2, Appli |
| 34 | 31 | 55.4 | 866 | 4 | US-09-621-855-2 | Sequence 2, Appli |
| 35 | 31 | 55.4 | 1026 | 1 | US-07-998-003A-95 | Sequence 95, Appl |
| 36 | 31 | 55.4 | 1026 | 1 | US-08-453-274B-95 | Sequence 95, Appl |
| 37 | 31 | 55.4 | 1026 | 1 | US-08-453-695A-95 | Sequence 95, Appl |
| 38 | 31 | 55.4 | 1026 | 1 | US-08-268-161A-95 | Sequence 95, Appl |
| 39 | 31 | 55.4 | 1026 | 2 | US-08-453-702A-95 | Sequence 95, Appl |
| 40 | 31 | 55.4 | 1026 | 4 | US-09-099-639-95 | Sequence 95, Appl |
| 41 | 31 | 55.4 | 1026 | 5 | PCT-US93-12588-95 | Sequence 95, Appl |
| 42 | 31 | 55.4 | 1026 | 5 | PCT-US95-08071-95 | Sequence 95, Appl |
| 43 | 31 | 55.4 | 1203 | 1 | US-07-998-003A-103 | Sequence 103, App |
| 44 | 31 | 55.4 | 1203 | 1 | US-08-453-274B-103 | Sequence 103, App |
| 45 | 31 | 55.4 | 1203 | 1 | US-08-453-695A-103 | Sequence 103, App |

ALIGNMENTS

```
RESULT   1
US-08-676-242-15
; Sequence 15, Application US/08676242C
; Patent No. 6143719
; GENERAL INFORMATION:
;  APPLICANT: The Regents of the University of Michigan
;  APPLICANT: Schmaier, Alvin H.
;  APPLICANT: Hasan, Ahmed A.K.
;  TITLE OF INVENTION: Bradykinin Analogs As Selective Thrombin Inhibitors
;  FILE REFERENCE: 8820-2 US
;  CURRENT APPLICATION NUMBER: US/08/676,242C
;  CURRENT FILING DATE: 2000-07-16
;  EARLIER APPLICATION NUMBER: 60/000,096
;  EARLIER FILING DATE: 1995-06-09
;  EARLIER APPLICATION NUMBER: PCT/US96/09940
;  EARLIER FILING DATE: 1996-06-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 26
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Bradykinin
;   OTHER INFORMATION: analog
US-08-676-242-15


  Query Match             100.0%;  Score 56;  DB 4;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 0.00029;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           ||||||||||||
Db       8 IDNVKKARVQVV 19


RESULT   2
US-08-193-114B-1
; Sequence 1, Application US/08193114B
; Patent No. 5472945
;  GENERAL INFORMATION:
;    APPLICANT:  Schmaier, Alvin H.
;    APPLICANT:  Jiang, Yongping
;    TITLE OF INVENTION:  Modulation of Blood
;    TITLE OF INVENTION:  Pressure and Inhibition of Platelet Activation
;    TITLE OF INVENTION:  with Kininogen Fragment
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seidel, Gonda, Lavorgna &
```

```
;       ADDRESSEE:  Monaco, P.C.
;       STREET:  1800 Two Penn Center Plaza
;       CITY:  Philadelphia
;       STATE:  Pennsylvania
;       COUNTRY:  U.S.A.
;       ZIP:  19102
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;       COMPUTER:  IBM PS/2
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WordPerfect 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/193,114B
;       FILING DATE:  9 February 1994
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  U.S. Application
;       APPLICATION NUMBER:  Serial No. 5472945 07/744,545
;       FILING DATE:  13 August 1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Monaco, Daniel A.
;       REGISTRATION NUMBER:  30,480
;       REFERENCE/DOCKET NUMBER:  6056-137 CIl
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (215) 568-8383
;       TELEFAX:  (215) 568-5549
;       TELEX:  No. 5472945e
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  117 amino acids
;       TYPE:  peptide
;       TOPOLOGY:  linear
US-08-193-114B-1


  Query Match             100.0%;  Score 56;  DB 1;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 0.0014;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           ||||||||||||
Db      38 IDNVKKARVQVV 49


RESULT   3
PCT-US92-06809-1
; Sequence 1, Application PC/TUS9206809
;  GENERAL INFORMATION:
;     APPLICANT:  Schmaier, Alvin H.
;     APPLICANT:  Jiang, Yongping
;     TITLE OF INVENTION:  Modulation of Blood
;     TITLE OF INVENTION:  Pressure by Altering Bradykinin Levels
;     NUMBER OF SEQUENCES:  2
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Temple University - Of the
;       ADDRESSEE:  Commonwealth System of Higher Education
;       STREET:  406 University Services
;       STREET:  Building
;       CITY:  Philadelphia
;       STATE:  Pennsylvania
;       COUNTRY:  U.S.A.
;       ZIP:  19122
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;       COMPUTER:  IBM PS/2
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WordPerfect 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US92/06809
;       FILING DATE:  19910813
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  U.S. Application
;       APPLICATION NUMBER:  Serial No.  744,545
;       FILING DATE:  13 August 1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Monaco, Daniel A.
;       REGISTRATION NUMBER:  30,480
;       REFERENCE/DOCKET NUMBER:  6056-137
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (215) 568-8383
;       TELEFAX:  (215) 568-5549
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  117 amino acids
;       TYPE:  AMINO ACID
;       TOPOLOGY:  linear
PCT-US92-06809-1


  Query Match             100.0%;  Score 56;  DB 5;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 0.0014;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           ||||||||||||
Db      38 IDNVKKARVQVV 49


RESULT   4
US-09-359-161-7
; Sequence 7, Application US/09359161A
; Patent No. 6342656
; GENERAL INFORMATION:
;  APPLICANT: Bradford, Kent J.
;  APPLICANT: Dahal, Peetambar
;  APPLICANT: Yang, Hong
;  APPLICANT: Cooley, Michael
;  APPLICANT: Downie, Bruce
;  APPLICANT: Gee, Oliver
;  APPLICANT: The Regents of the University of California
;  TITLE OF INVENTION: Regulation of Source-Sink Relationships and Responses
;  TITLE OF INVENTION: to Stress Conditions in Plants
;  FILE REFERENCE: 023070-095900US
;  CURRENT APPLICATION NUMBER: US/09/359,161A
;  CURRENT FILING DATE: 1999-07-21
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
;   FEATURE:
;   OTHER INFORMATION: yeast sucrose nonfermenting protein kinase 1
;   OTHER INFORMATION: kinase subunit (SNF1)
US-09-359-161-7


  Query Match             64.3%;  Score 36;  DB 4;  Length 322;
  Best Local Similarity   66.7%;  Pred. No. 22;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARV 9
           :||::||||
Db     284 MDNIRKARV 292


RESULT   5
US-08-441-139-5
; Sequence 5, Application US/08441139
; Patent No. 5773245
;  GENERAL INFORMATION:
;    APPLICANT:  Wittrup, Dr. Karl D.
;    APPLICANT:  Robinson, Anne S.
;    TITLE OF INVENTION:  METHODS FOR INCREASING SECRETION OF
```

```
;    TITLE OF INVENTION:  RECOMBINANTLY EXPRESSED PROTEINS
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCULLY, SCOTT, MURPHY & PRESSER
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/441,139
;      FILING DATE:  15-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/089,997
;      FILING DATE:  06-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8646
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  516-742-4343
;      TELEFAX:  516-742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  679 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-441-139-5


  Query Match             60.7%;  Score 34;  DB 1;  Length 679;
  Best Local Similarity   54.5%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQV 11
           :|| |  |:||
Db     476 VDNQKAVRIQV 486


RESULT   6
US-09-537-682-1
; Sequence 1, Application US/09537682
; Patent No. 6303357
; GENERAL INFORMATION:
;  APPLICANT: TAKEUCHI, Kenichi
;  APPLICANT: KOIDE, Yoshinao
;  APPLICANT: NAKANISHI, Yuji
;  APPLICANT: SUZUKI, Satoru
;  TITLE OF INVENTION: L-ALPHA-GLYCEROPHOSPHATE OXIDASE GENE, RECOMBINATNT
;  TITLE OF INVENTION: DNA, AND METHOD FOR PRODUCING MODIFIED
;  TITLE OF INVENTION: L-ALPHA-GLYCEROPHOSPHATE OXIDASE GENE
;  FILE REFERENCE: A20-121814C/KI
;  CURRENT APPLICATION NUMBER: US/09/537,682
;  CURRENT FILING DATE: 2000-03-29
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 607
;   TYPE: PRT
;   ORGANISM: Enterococcus faecium No. 6303357 7044
US-09-537-682-1


  Query Match             58.9%;  Score 33;  DB 4;  Length 607;
  Best Local Similarity   85.7%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 IDNVKKA 7
           |||:|||
Db     178 IDNIKKA 184


RESULT   7
US-09-188-930-174
; Sequence 174, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Strachan, Lorna
;  APPLICANT: Sleeman, Matthew
;  APPLICANT: Onrust, Rene
;  APPLICANT: Murison, James Greg
;  TITLE OF INVENTION: Compositions Isolated From Skin Cells
;  TITLE OF INVENTION: and Methods For Their Use
;  FILE REFERENCE: 11000.1011c1
;  CURRENT APPLICATION NUMBER: US/09/188,930A
;  CURRENT FILING DATE: 1998-11-09
;  NUMBER OF SEQ ID NOS: 348
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 174
;   LENGTH: 137
;   TYPE: PRT
;   ORGANISM: Human
US-09-188-930-174


  Query Match             57.1%;  Score 32;  DB 4;  Length 137;
  Best Local Similarity   75.0%;  Pred. No. 50;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 DNVKKARV 9
           | |||||:
Db      51 DKVKKARI 58


RESULT   8
US-08-622-277A-14
; Sequence 14, Application US/08622277A
; Patent No. 6001580
;  GENERAL INFORMATION:
;    APPLICANT:  Tani, Akiyoshi
;    APPLICANT:  Ichimori, Yuzo
;    TITLE OF INVENTION:  Method For Assaying MAP Kinase
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W., Suite 700
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/622,277A
;      FILING DATE:  27-MAR-1996
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 320577-1995
;      FILING DATE:  08-DEC-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 305456-1995
;      FILING DATE:  24-NOV-1995
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 070125-1995
;      FILING DATE:  28-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toohey, Kimberlin M.
;      REGISTRATION NUMBER:  35,391
;      REFERENCE/DOCKET NUMBER:  04221.0039-00000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)408-4000
;      TELEFAX:  (202)408-4400
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  360 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-622-277A-14


  Query Match             57.1%;  Score 32;  DB 3;  Length 360;
  Best Local Similarity   60.0%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 DNVKKARVQV 11
           ||| | || :
Db      44 DNVNKVRVAI 53


RESULT   9
US-09-025-580-2
; Sequence 2, Application US/09025580
; Patent No. 6162613
;  GENERAL INFORMATION:
;    APPLICANT:  Su, Michael Shin-San
;    APPLICANT:  Fox, Ted
;    APPLICANT:  Wilson, Keith Phillip
;    APPLICANT:  Germann, Ursula A.
;    TITLE OF INVENTION:  Methods For Designing Inhibitors of
;    TITLE OF INVENTION:  Serine/Thereonine Kinases and Tyrosine Kinase
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Neave
;      STREET:  1251 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  10020
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/025,580
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haley, James F.
;      REGISTRATION NUMBER:  27,794
;      REFERENCE/DOCKET NUMBER:  VPI 97-104
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 596-9000
;      TELEFAX:  (212) 596-9090
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  360 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  103..104
;      OTHER INFORMATION:  /note= "amino acid 103 is
;      OTHER INFORMATION:  isoleucine or leucine"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  105..106
;      OTHER INFORMATION:  /product= "OTHER"
;      OTHER INFORMATION:  /note= "amino acid 105 is glutamine, threonine or alanine"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  106..107
;      OTHER INFORMATION:  /product= "OTHER"
;      OTHER INFORMATION:  /note= "amino acid 106 is aspartic acid or histidine"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  109..110
;      OTHER INFORMATION:  /product= "OTHER"
;      OTHER INFORMATION:  /note= "amino acid 109 is glutamic acid or glycine"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  110..111
;      OTHER INFORMATION:  /product= "OTHER"
;      OTHER INFORMATION:  /note= "amino acid 110 is threonine or alanine"
US-09-025-580-2


  Query Match             57.1%;  Score 32;  DB 4;  Length 360;
  Best Local Similarity   60.0%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 DNVKKARVQV 11
           ||| | || :
Db      44 DNVNKVRVAI 53


RESULT  10
US-08-459-953A-9
; Sequence 9, Application US/08459953A
; Patent No. 6030822
;  GENERAL INFORMATION:
;    APPLICANT:  Lechner, Cornelia
;    APPLICANT:  Moller, Niels P.H.
;    APPLICANT:  Ullrich, Axel
;    TITLE OF INVENTION:  EXTRACELLULAR SIGNAL-RELATED
;    TITLE OF INVENTION:  KINASE, SEQUENCES, AND
;    TITLE OF INVENTION:  METHODS OF PRODUCTION
;    TITLE OF INVENTION:  AND USE
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  Microsoft Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/459,953A
;      FILING DATE:  June 2, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/029,494
;      FILING DATE:  March 19, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  209/267
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  9 :
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  380 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-459-953A-9


  Query Match             57.1%;  Score 32;  DB 3;  Length 380;
  Best Local Similarity   60.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 DNVKKARVQV 11
           ||| | || :
Db      64 DNVNKVRVAI 73


RESULT  11
US-08-979-917A-2
; Sequence 2, Application US/08979917A
; Patent No. 6118050
;  GENERAL INFORMATION:
;    APPLICANT:  STURNER, STEPHEN
;    APPLICANT:  HIRAYAMA, LYNNE MIYO
;    APPLICANT:  SINGH, BIJAY
;    APPLICANT:  BASCOMB, NEWELL
;    TITLE OF INVENTION:  HPPD GENE AND INHIBITORS
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Darby & Darby PC
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,917A
;      FILING DATE:  25-JUL-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/022,604
;      FILING DATE:  25-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zitron, Anne E.
;      REGISTRATION NUMBER:  41,391
;      REFERENCE/DOCKET NUMBER:  0646/1B917-US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-527-7700
;      TELEFAX:  212-753-6237
;      TELEX:  236687
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  392 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6118050e
US-08-979-917A-2


  Query Match             57.1%;  Score 32;  DB 3;  Length 392;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 DNVKKARVQV 11
           :|:| |::||
Db     300 ENLKSAKIQV 309


RESULT  12
US-09-252-292C-23
; Sequence 23, Application US/09252292C
; Patent No. 6245968
; GENERAL INFORMATION:
;  APPLICANT:  Boudec, Phillipe
;  APPLICANT:  Rodgers, Matthew
;  APPLICANT:  Dumas, Florence
;  TITLE OF INVENTION:  Mutated  hydroxyphenylpyruvate dioxygenase, DNA
;  TITLE OF INVENTION:  sequence and isolation of plants which contain such a
;  TITLE OF INVENTION:  gene and which are tolerant to herbicides
;  FILE REFERENCE:  5500*31
;  CURRENT APPLICATION NUMBER: US/09/252,292C
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 08/982,772
;  PRIOR FILING DATE:  1997-12-02
;  NUMBER OF SEQ ID NOS:  31
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 23
;   LENGTH: 393
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-252-292C-23


  Query Match             57.1%;  Score 32;  DB 4;  Length 393;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 DNVKKARVQV 11
           :|:| |::||
Db     301 ENLKSAKIQV 310


RESULT  13
US-08-846-762-74
; Sequence 74, Application US/08846762A
; Patent No. 5994072
; GENERAL INFORMATION:
;  APPLICANT: Lam, Joseph S.
;  APPLICANT: Burrows, Lori
;  APPLICANT: Charter, Deborah
;  APPLICANT: de Kievit, Teresa
;  TITLE OF INVENTION: No. 5994072el Proteins Involved in the Synthesis and Assembly
;  TITLE OF INVENTION: of O-Antigen in Pseudomonas Aeruginosa
;  FILE REFERENCE: 6580-089
;  CURRENT APPLICATION NUMBER: US/08/846,762A
;  CURRENT FILING DATE: 1997-04-30
;  NUMBER OF SEQ ID NOS: 100
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 74
;   LENGTH: 423
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-08-846-762-74


  Query Match             57.1%;  Score 32;  DB 2;  Length 423;
  Best Local Similarity   33.3%;  Pred. No. 1.6e+02;
  Matches    4;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           :| : :||: :|
Db      45 VDTINQARIHIV 56


RESULT  14
```

```
US-09-342-648-10
; Sequence 10, Application US/09342648
; Patent No. 6248584
; GENERAL INFORMATION:
;  APPLICANT: Cahoon, Rebecca E.
;  APPLICANT: Odell, Joan
;  APPLICANT: Rafalski, Antoni
;  TITLE OF INVENTION: Transcription Coactivators
;  FILE REFERENCE: BB-1169-B
;  CURRENT APPLICATION NUMBER: US/09/342,648
;  CURRENT FILING DATE: 1999-06-29
;  EARLIER APPLICATION NUMBER: 60/092,659
;  EARLIER FILING DATE: July 13, 1998
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
;   LENGTH: 890
;   TYPE: PRT
;   ORGANISM: Ajellomyces capsulatus
US-09-342-648-10


  Query Match           57.1%;  Score 32;  DB 4;  Length 890;
  Best Local Similarity  55.6%;  Pred. No. 3.5e+02;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 DNVKKARVQ 10
           :||:||::|
Db     475 ENVQKAKIQ 483


RESULT  15
US-08-750-717-2
; Sequence 2, Application US/08750717
; Patent No. 6180109
;  GENERAL INFORMATION:
;    APPLICANT:  MOORMANN, Robertus J. M.
;    APPLICANT:  VAN RIJN, Petrus A.
;    TITLE OF INVENTION:  Nucleotide Sequences of Pestivirus
;    TITLE OF INVENTION:  Strains, Polypeptides Encoded by These Sequences and Use
;    TITLE OF INVENTION:  Thereof for Diagnosis and Prevention of Pestivirus
;    TITLE OF INVENTION:  Infections
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/750,717
;      FILING DATE:  24-DEC-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 94201743.5
;      FILING DATE:  17-JUN-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/NL95/00214
;      FILING DATE:  16-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  BO 39123
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-521-2297
;      TELEFAX:  703-685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3898 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-750-717-2


  Query Match           57.1%;  Score 32;  DB 4;  Length 3898;
  Best Local Similarity  75.0%;  Pred. No. 1.6e+03;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 IDNVKKAR 8
             |||:|| |
Db    3425 IDNLKKGR 3432


Search completed: July  1, 2002, 16:18:01
Job time: 45 sec
```

OM protein - protein search, using sw model

Run on: July 1, 2002, 16:20:38 ; Search time 46.58 Seconds
(without alignments)
66.012 Million cell updates/sec

Title: US-09-461-061A-4
Perfect score: 161
Sequence: 1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 161 | 100.0 | 427 | 1 | KGHUL1 | kininogen, LMW pre |
| 2 | 161 | 100.0 | 644 | 1 | KGHUH1 | kininogen, HMW pre |
| 3 | 116 | 72.0 | 434 | 1 | KGBOL2 | kininogen, LMW II |
| 4 | 116 | 72.0 | 619 | 1 | KGBOH2 | kininogen, HMW II |
| 5 | 112 | 69.6 | 433 | 2 | A28055 | K-kininogen, LMW I |
| 6 | 112 | 69.6 | 436 | 1 | KGBOL1 | kininogen, LMW I p |
| 7 | 112 | 69.6 | 621 | 1 | KGBOH1 | kininogen, HMW I p |
| 8 | 112 | 69.6 | 639 | 2 | A25486 | kininogen, HMW I p |
| 9 | 99 | 61.5 | 423 | 1 | KGRTM | major acute phase |
| 10 | 98 | 60.9 | 430 | 1 | KGRTT1 | T-kininogen I prec |
| 11 | 98 | 60.9 | 430 | 2 | A23897 | major acute phase |
| 12 | 98 | 60.9 | 430 | 2 | B28055 | T-kininogen, LMW I |
| 13 | 52.5 | 32.6 | 498 | 2 | S49776 | hypothetical prote |
| 14 | 52 | 32.3 | 217 | 2 | B97948 | hypothetical prote |
| 15 | 52 | 32.3 | 701 | 2 | A44943 | heat shock protein |
| 16 | 51.5 | 32.0 | 803 | 2 | F90485 | hypothetical prote |
| 17 | 51 | 31.7 | 139 | 2 | T33740 | hypothetical prote |
| 18 | 51 | 31.7 | 209 | 2 | T10057 | cysteine proteinas |
| 19 | 51 | 31.7 | 364 | 2 | AH0548 | probable lipoprote |
| 20 | 51 | 31.7 | 523 | 2 | A97177 | site-specific reco |
| 21 | 50 | 31.1 | 306 | 2 | T38834 | hypothetical coile |
| 22 | 49.5 | 30.7 | 375 | 2 | A32827 | fetuin precursor - |
| 23 | 49.5 | 30.7 | 709 | 2 | G96610 | probable disease r |
| 24 | 49 | 30.4 | 313 | 2 | E81356 | hypothetical prote |
| 25 | 49 | 30.4 | 459 | 2 | G90558 | aminopeptidase (le |
| 26 | 49 | 30.4 | 768 | 2 | AH1085 | hypothetical prote |
| 27 | 49 | 30.4 | 1034 | 2 | G90591 | hypothetical prote |
| 28 | 49 | 30.4 | 5005 | 2 | F82884 | hypothetical prote |
| 29 | 48 | 29.8 | 377 | 2 | G70143 | hypothetical prote |
| 30 | 48 | 29.8 | 624 | 2 | T28423 | ORF MSV261 leucine |
| 31 | 48 | 29.8 | 1231 | 2 | A86359 | hypothetical prote |
| 32 | 47.5 | 29.5 | 107 | 2 | T21642 | hypothetical prote |
| 33 | 47.5 | 29.5 | 107 | 2 | T24886 | hypothetical prote |
| 34 | 47.5 | 29.5 | 392 | 2 | S50139 | translation elonga |
| 35 | 47.5 | 29.5 | 452 | 2 | S68466 | translation elonga |
| 36 | 47.5 | 29.5 | 452 | 2 | S62768 | translation elonga |
| 37 | 47.5 | 29.5 | 452 | 2 | S62767 | translation elonga |
| 38 | 47.5 | 29.5 | 454 | 2 | G75105 | hypothetical prote |
| 39 | 47.5 | 29.5 | 645 | 2 | S03208 | type III site-spec |
| 40 | 47.5 | 29.5 | 890 | 2 | A97750 | DNA mismatch repai |
| 41 | 47.5 | 29.5 | 939 | 2 | E90558 | topoisomerase iv s |
| 42 | 47 | 29.2 | 111 | 2 | A28793 | cystatin - puff ad |
| 43 | 47 | 29.2 | 224 | 2 | C82903 | cytidylate kinase |
| 44 | 47 | 29.2 | 256 | 2 | E71807 | hypothetical prote |
| 45 | 47 | 29.2 | 431 | 2 | T29188 | hypothetical prote |

ALIGNMENTS

RESULT 1
KGHUL1
kininogen, LMW precursor [validated] - human
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Homo sapiens (man)
C;Date: 06-Jul-1982 #sequence_revision 27-Nov-1985 #text_change 08-Dec-2000
C;Accession: A01280; B25276; A27900; A27699; A31905; A34030
R;Ohkubo, I.; Kurachi, K.; Takasawa, T.; Shiokawa, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A;Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its ide
A;Reference number: A90490; MUID:85122621
A;Accession: A01280
A;Molecule type: mRNA
A;Residues: 1-427 <OHK>
A;Cross-references: GB:K02566; NID:g177889; PIDN:AAA35497.1; PID:g177890
R;Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A;Title: Cloning and sequence analysis of cDNAs for human high molecular weight and l
A;Reference number: A92544; MUID:85234582
A;Accession: B25276
A;Molecule type: mRNA
A;Residues: 1-427 <TAK>
A;Cross-references: GB:M11437; NID:g186751; PIDN:AAB59551.1; PID:g386853
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Rauth, G.; Mueller-Esterl, W.
in Kinins IV, part A, Greenbaum, L.M., and Margolius, H.S., eds., pp.91-95, Plenum, N
A;Title: Amino acid sequence of the light chain of human low molecular mass kininogen
A;Reference number: A27900
A;Accession: A27900
A;Molecule type: protein
A;Residues: 390-427 <LOT>
R;Mindroiu, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
Biochem. Biophys. Res. Commun. 152, 519-526, 1988
A;Title: A new kinin moiety in human plasma kininogens.
A;Reference number: A27699; MUID:88209021
A;Accession: A27699
A;Molecule type: protein
A;Residues: 380-389 <MIN>
R;Maeda, H.; Matsumura, Y.; Kato, H.
J. Biol. Chem. 263, 16051-16054, 1988
A;Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic f
A;Reference number: A31905; MUID:89034061
A;Accession: A31905
A;Molecule type: protein
A;Residues: 381-389 <MAE>
R;Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
Biochem. Biophys. Res. Commun. 150, 511-516, 1988
A;Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human p
A;Reference number: A34030; MUID:88106632
A;Accession: A34030
A;Molecule type: protein
A;Residues: 380-389 <SAS>

R;Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 260, 8610-8617, 1985
A;Title: Structural organization of the human kininogen gene and a model for its evoluti
A;Reference number: A92545; MUID:85234583
A;Contents: annotation; gene organization
R;Pierce, J.V.
Fed. Proc. 27, 52-57, 1968
A;Title: Structural features of plasma kinins and kininogens.
A;Reference number: A91455; MUID:90255622
A;Contents: annotation; bradykinin
C;Comment: The LMW kininogen precursor is produced from the same gene as the HMW form (s
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Genetics:
A;Gene: GDB:KNG
A;Cross-references: GDB:125256; OMIM:228960
A;Map position: 3q27-3q27
A;Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3; 401/3
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; glyc
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-427/Product: LMW prokininogen (kininogen I) #status predicted <MAT>
F;19-389,390-427/Product: LMW kininogen II #status predicted <MAT2>
F;19-379/Product: LMW kininogen heavy chain #status predicted <HCH>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>
F;380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
F;390-427/Product: LMW kininogen light chain #status experimental <LCH>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;28-407,83-94,107-126,142-145,206-218,229-248,264-267,328-340,351-370/Disulfide bonds:
F;48,169,205,294/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
F;383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F;389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;401/Binding site: carbohydrate (Thr) (covalent) #status absent

Query Match 100.0%; Score 161; DB 1; Length 427;
Best Local Similarity 100.0%; Pred. No. 6.8e-14;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32
||||||||||||||||||||||||||||||||
Db 281 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

RESULT 2
KGHUH1
kininogen, HMW precursor [validated] - human
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen; prokininogen
N;Contains: bradykinin (kallidin I); HMW kininogen I; HMW kininogen II; low molecular we
C;Species: Homo sapiens (man)
C;Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 08-Dec-2000
C;Accession: A01279; A25276; S32422; A91153; A24871; A27899; A27699; A31905; A34030; S02
R;Ohkubo, I.; Kurachi, K.; Takasawa, T.; Shiokawa, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A;Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its identi
A;Reference number: A90490; MUID:85122621
A;Accession: A01279
A;Molecule type: mRNA
A;Residues: 1-389 <OHK>
A;Cross-references: GB:K02566; NID:g177889
R;Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A;Title: Cloning and sequence analysis of cDNAs for human high molecular weight and low
A;Reference number: A92544; MUID:85234582
A;Accession: A25276
A;Molecule type: mRNA
A;Residues: 1-592,'I',594-644 <TAK>
A;Cross-references: GB:M11437; NID:g186751; PIDN:AAB59550.1; PID:g386852
R;Auerswald, E.A.; Roessler, D.; Mentele, R.; Assfalg-Machleidt, I.
FEBS Lett. 321, 93-97, 1993
A;Title: Cloning, expression and characterization of human kininogen domain 3.
A;Reference number: S32422; MUID:93223854
A;Accession: S32422
A;Molecule type: mRNA
A;Residues: 'ANSM',253-377 <AUE>
A;Note: differences are due to known cloning artifacts
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A;Title: The amino acid sequence of the light chain of human high-molecular-mass kini
A;Reference number: A91153; MUID:86030270
A;Accession: A91153
A;Molecule type: protein
A;Residues: 379-644 <LOT>
A;Note: the bradykinin sequence preceding the light chain sequence was not determined
R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
Eur. J. Biochem. 154, 471-478, 1986
A;Title: Completion of the primary structure of human high-molecular-mass kininogen.
A;Reference number: A24871; MUID:86108361
A;Accession: A24871
A;Molecule type: protein
A;Residues: 'Z',20-380 <KEL1>
R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
in Kinins IV, Greenbaum, L.M., and Margolius, H.S., ed., pp.85-89, Plenum Press, New
A;Title: Amino acid sequence of the light chain of human high molecular mass kininoge
A;Reference number: A27899
A;Accession: A27899
A;Molecule type: protein
A;Residues: 379-389,'K',390-407,'Q',409-644 <KEL2>
R;Mindroiu, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
Biochem. Biophys. Res. Commun. 152, 519-526, 1988
A;Title: A new kinin moiety in human plasma kininogens.
A;Reference number: A27699; MUID:88209021
A;Accession: A27699
A;Molecule type: protein
A;Residues: 380-389 <MIN>
R;Maeda, H.; Matsumura, Y.; Kato, H.
J. Biol. Chem. 263, 16051-16054, 1988
A;Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic f
A;Reference number: A31905; MUID:89034061
A;Accession: A31905
A;Molecule type: protein
A;Residues: 381-389 <MAE>
R;Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
Biochem. Biophys. Res. Commun. 150, 511-516, 1988
A;Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human p
A;Reference number: A34030; MUID:88106632
A;Accession: A34030
A;Molecule type: protein
A;Residues: 380-389 <SAS>
R;Lenarcic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
Biol. Chem. Hoppe-Seyler 369, 257-261, 1988
A;Title: Human cathepsin B and cysteine proteinase inhibitors (CPIs) in inflammatory
A;Reference number: S02482; MUID:89076517
A;Accession: S02482
A;Molecule type: protein
A;Residues: 1-19;189-192;310-314;381-389 <LEN1>
R;Kato, H.; Matsumura, Y.; Maeda, H.
FEBS Lett. 232, 252-254, 1988
A;Title: Isolation and identification of hydroxyproline analogues of bradykinin in hu
A;Reference number: A61495; MUID:88211869
A;Accession: A61495
A;Molecule type: protein
A;Residues: 380-389 <KAT1>
A;Experimental source: urine
A;Note: this peptide had Pro-383 modified to 4-hydroxyproline
A;Accession: B61495
A;Molecule type: protein
A;Residues: 381-389 <KAT2>
A;Experimental source: urine
A;Note: this peptide had Pro-383 modified to 4-hydroxyproline
A;Accession: C61495

A;Molecule type: protein
A;Residues: 380-389 <KAT3>
R;Lenarcic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
FEBS Lett. 280, 211-215, 1991
A;Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
A;Reference number: S14303; MUID:91192133
A;Accession: S14447
A;Molecule type: protein
A;Residues: 264-359,'N',361-375 <LEN2>
R;Little, S.S.; Johnson, D.A.
Biochem. J. 307, 341-346, 1995
A;Title: Human mast cell tryptase isoforms: separation and examination of substrate-spec
A;Reference number: S55239; MUID:95251593
A;Accession: S55239
A;Molecule type: protein
A;Residues: 450-452,'X',454,'X',456 <LIT>
R;Straczek, J.; Maachi, F.; le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabet, P.; Bellevil
FEBS Lett. 373, 207-211, 1995
A;Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like
A;Reference number: S68059; MUID:96033974
A;Accession: S68059
A;Molecule type: protein
A;Residues: 431-434 <STR>
R;Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 260, 8610-8617, 1985
A;Title: Structural organization of the human kininogen gene and a model for its evoluti
A;Reference number: A92545; MUID:85234583
A;Contents: annotation; gene organization
R;Pierce, J.V.
Fed. Proc. 27, 52-57, 1968
A;Title: Structural features of plasma kinins and kininogens.
A;Reference number: A91455; MUID:90255622
A;Contents: annotation; bradykinin
C;Comment: The HMW kininogen precursor and the LMW form are produced from the same gene
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Genetics:
A;Gene: GDB:KNG
A;Cross-references: GDB:125256; OMIM:228960
A;Map position: 3q27-3q27
A;Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>
F;19-379,390-644/Product: HMW kininogen II #status experimental <MAT2>
F;19-379/Domain: HMW kininogen heavy chain #status experimental <HCH>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>
F;380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
F;390-644/Domain: HMW kininogen light chain #status experimental <LCH>
F;421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
F;431-434/Product: low molecular weight growth promoting factor #status experimental <GF
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimer
F;28-614,83-94,107-126,142-145,206-218,229-248,264-267,328-340,351-370/Disulfide bonds:
F;48/Binding site: carbohydrate (Asn) (covalent) #status absent
F;169,205,294/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
F;383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F;389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;401,533,542,546,557,571,593,628/Binding site: carbohydrate (Thr) (covalent) #status ex
F;577/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 161; DB 1; Length 644;
Best Local Similarity 100.0%; Pred. No. 1e-13;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32
||||||||||||||||||||||||||||||||
Db 281 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

RESULT 3
KGBOL2
kininogen, LMW II precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 28-May-1999
C;Accession: A01284
R;Nawa, H.; Kitamura, N.; Hirose, T.; Asai, M.; Inayama, S.; Nakanishi, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 90-94, 1983
A;Title: Primary structures of bovine liver low molecular weight kininogen precursors
A;Reference number: A93984; MUID:83117859
A;Accession: A01284
A;Molecule type: mRNA
A;Residues: 1-434 <NAW>
A;Cross-references: GB:V00427; GB:J00011; NID:g489; PIDN:CAA23710.1; PID:g490
C;Comment: The LMW kininogen precursor is produced from the same gene as the HMW form
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; g
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-434/Product: LMW kininogen II #status predicted <MAT>
F;19-377/Product: LMW kininogen I heavy chain #status predicted <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;261-372/Domain: cystatin homology <CY3>
F;377-386/Product: lysyl-bradykinin (kallidin II) #status predicted <KBDY>
F;378-386/Product: bradykinin (kallidin I) #status predicted <BDY>
F;387-434/Product: LMW kininogen I light chain #status experimental <LCH>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F;27-404,82-93,106-125,141-144,205-217,228-247,261-264,325-337,348-367/Disulfide bond
F;47,87,168,169,197,204,280/Binding site: carbohydrate (Asn) (covalent) #status predi
F;376-377/Cleavage site: Met-Lys (kallikrein) #status predicted
F;380/Modified site: 4-hydroxyproline (Pro) #status predicted
F;386-387/Cleavage site: Arg-Ser (kallikrein) #status predicted

Query Match 72.0%; Score 116; DB 1; Length 434;
Best Local Similarity 74.2%; Pred. No. 7.1e-08;
Matches 23; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
| |:| |||||:: ||||||| |||| ||||
Db 279 LNHSIAKLNAEHDGTFYFKIDTVKKATVQVV 309

RESULT 4
KGBOH2
kininogen, HMW II precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
C;Accession: A01282; A91923; A91941; A91938; B29559
R;Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
Nature 305, 545-549, 1983
A;Title: A single gene for bovine high molecular weight and low molecular weight kini
A;Reference number: A93317; MUID:84014106
A;Accession: A01282
A;Molecule type: mRNA
A;Residues: 1-619 <KIT>
A;Cross-references: GB:V01492; GB:K01758; NID:g493; PIDN:CAA24736.1; PID:g494
R;Kato, H.; Nagasawa, S.; Suzuki, T.
J. Biochem. 67, 313-323, 1970
A;Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds a
A;Reference number: A91923; MUID:70180420
A;Accession: A91923

A;Molecule type: protein
A;Residues: 376-391 <KAT>
R;Han, Y.N.; Kato, H.; Iwanaga, S.; Suzuki, T.
J. Biochem. 79, 1201-1222, 1976
A;Title: Primary structure of bovine plasma high-molecular-weight kininogen. The amino a
A;Reference number: A91941; MUID:76260155
A;Accession: A91941
A;Molecule type: protein
A;Residues: 387-455 <HAN>
A;Note: 398-Pro, 401-Val, and 455-Lys were also found
R;Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
J. Biochem. 77, 55-68, 1975
A;Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Ami
A;Reference number: A91938; MUID:75170265
A;Accession: A91938
A;Molecule type: protein
A;Residues: 456-496 <HA2>
R;Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga,
J. Biol. Chem. 262, 2768-2779, 1987
A;Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of c
A;Reference number: A92627; MUID:87137530
A;Accession: B29559
A;Molecule type: protein
A;Residues: 'Z',20-104,'E',106-256,'XX',257-376 <SUE>
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A;Title: The amino acid sequence of the light chain of human high-molecular-mass kininog
A;Reference number: A91153; MUID:86030270
A;Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
R;Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
Seikagaku 56, 808, 1984
A;Title: Disulfide bonds in bovine HMW kininogens.
A;Reference number: A94300
A;Contents: annotation; disulfide bonds
A;Note: article in Japanese
C;Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-619/Product: HMW kininogen II #status predicted <MAT>
F;19-376/Product: HMW kininogen II heavy chain #status experimental <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;261-372/Domain: cystatin homology <CY3>
F;377-386/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;378-386/Product: bradykinin (kallidin I) #status experimental <BDY>
F;387-619/Product: HMW kininogen II light chain #status experimental <LCH>
F;418-488/Region: glycine/histidine/lysine-rich
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;27-589,82-93,106-125,141-144,205-217,228-247,261-264,325-337,348-367/Disulfide bonds:
F;47/Binding site: carbohydrate (Asn) (covalent) #status absent
F;87,168,169,204,280/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
F;197/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F;376-377/Cleavage site: Met-Lys (kallikrein) #status experimental
F;380/Modified site: 4-hydroxyproline (Pro) #status predicted
F;386-387/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;396,400,404,510/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;397,398,518,522,534,546,551,568/Binding site: carbohydrate (Thr) (covalent) #status ex
F;496-497/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match 72.0%; Score 116; DB 1; Length 619;
Best Local Similarity 74.2%; Pred. No. 1e-07;
Matches 23; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
| |:| |||||:: ||||||| |||| ||||
Db 279 LNHSIAKLNAEHDGTFYFKIDTVKKATVQVV 309

RESULT 5
A28055
K-kininogen, LMW I precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 15-Nov-1996
C;Accession: A28055
R;Furuto-Kato, S.; Matsumoto, A.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 12054-12059, 1985
A;Title: Primary structures of the mRNAs encoding the rat precursors for bradykinin a
nhibitor.
A;Reference number: A92496; MUID:86008264
A;Accession: A28055
A;Molecule type: mRNA
A;Residues: 1-433 <FUR>
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-433/Product: K-kininogen, LMW I #status predicted <MAT>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>

Query Match 69.6%; Score 112; DB 2; Length 433;
Best Local Similarity 74.2%; Pred. No. 2.4e-07;
Matches 23; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
| |:| :|||||| ||||||| |||| |||
Db 282 LGHSIAQLNAENNHTFYFKIDTVKKATSQVV 312

RESULT 6
KGBOL1
kininogen, LMW I precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
C;Accession: A01283
R;Nawa, H.; Kitamura, N.; Hirose, T.; Asai, M.; Inayama, S.; Nakanishi, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 90-94, 1983
A;Title: Primary structures of bovine liver low molecular weight kininogen precursors
A;Reference number: A93984; MUID:83117859
A;Accession: A01283
A;Molecule type: mRNA
A;Residues: 1-436 <NAW>
A;Cross-references: GB:J00010; GB:V00426; NID:g163256; PIDN:AAA30604.1; PID:g163257
C;Comment: The LMW kininogen precursor is produced from the same gene as the HMW form
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; g
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-436/Product: LMW kininogen I #status predicted <MAT>
F;19-378/Product: LMW kininogen I heavy chain #status predicted <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;379-388/Product: lysyl-bradykinin (kallidin II) #status predicted <KBDY>
F;380-388/Product: bradykinin (kallidin I) #status predicted <BDY>
F;389-436/Product: LMW kininogen I light chain #status experimental <LCH>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F;27-406,82-93,106-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bond
F;47,87,168,169,197,204/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;378-379/Cleavage site: Met-Lys (kallikrein) #status predicted
F;382/Modified site: 4-hydroxyproline (Pro) #status predicted
F;388-389/Cleavage site: Arg-Ser (kallikrein) #status predicted

```
  Query Match             69.6%;  Score 112;  DB 1;  Length 436;
  Best Local Similarity   71.0%;  Pred. No. 2.5e-07;
  Matches   22;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           |:|:| |||||::  |||||| |||| ||||
Db     281 LSHSIAKLNAEHDGAFYFKIDTVKKATVQVV 311


RESULT   7
KGBOH1
kininogen, HMW I precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
C;Accession: A01281; A91923; A91938; A29559
R;Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
Nature 305, 545-549, 1983
A;Title: A single gene for bovine high molecular weight and low molecular weight kininog
A;Reference number: A93317; MUID:84014106
A;Accession: A01281
A;Molecule type: mRNA
A;Residues: 1-621 <KIT>
A;Cross-references: GB:V01491; GB:K01757; NID:g491; PIDN:CAA24735.1; PID:g492
R;Kato, H.; Nagasawa, S.; Suzuki, T.
J. Biochem. 67, 313-323, 1970
A;Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds and
A;Reference number: A91923; MUID:70180420
A;Accession: A91923
A;Molecule type: protein
A;Residues: 378-393 <KAT>
R;Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
J. Biochem. 77, 55-68, 1975
A;Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Ami
A;Reference number: A91938; MUID:75170265
A;Accession: A91938
A;Molecule type: protein
A;Residues: 458-498 <HAN>
R;Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga,
J. Biol. Chem. 262, 2768-2779, 1987
A;Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of c
A;Reference number: A92627; MUID:87137530
A;Accession: A29559
A;Molecule type: protein
A;Residues: 'Z',20-123,'I',125-127,'I',129-378 <SUE>
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A;Title: The amino acid sequence of the light chain of human high-molecular-mass kininog
A;Reference number: A91153; MUID:86030270
A;Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
R;Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
Seikagaku 56, 808, 1984
A;Title: Disulfide bonds in bovine HMW kininogens.
A;Reference number: A94300
A;Contents: annotation; disulfide bonds
A;Note: article in Japanese
C;Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-621/Product: HMW prokininogen I #status predicted <MAT>
F;19-379/Product: HMW kininogen I heavy chain #status experimental <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;379-388/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;380-388/Product: bradykinin (kallidin I) #status experimental <BDY>
F;389-621/Product: HMW kininogen I light chain #status experimental <LCH>
F;417-488/Region: glycine/histidine/lysine-rich
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
F;27-591,82-93,106-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bond
F;87,168,169,204/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
F;197/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F;378-379/Cleavage site: Met-Lys (kallikrein) #status experimental
F;382/Modified site: 4-hydroxyproline (Pro) #status predicted
F;388-389/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;398,406,512/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;399,400,520,524,536,548,553,570/Binding site: carbohydrate (Thr) (covalent) #status
F;498-499/Cleavage site: Arg-Thr (kallikrein) #status experimental


  Query Match             69.6%;  Score 112;  DB 1;  Length 621;
  Best Local Similarity   71.0%;  Pred. No. 3.5e-07;
  Matches   22;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           |:|:| |||||::  |||||| |||| ||||
Db     281 LSHSIAKLNAEHDGAFYFKIDTVKKATVQVV 311


RESULT   8
A25486
kininogen, HMW I precursor - rat
N;Contains: bradykinin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Nov-1996
C;Accession: A25486
R;Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A;Title: Differing expression patterns and evolution of the rat kininogen gene family
A;Reference number: A92625; MUID:87137443
A;Accession: A25486
A;Molecule type: mRNA
A;Residues: 1-639 <KIT>
A;Note: the authors translated the codon CAA for residue 347 as Asn
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-639/Product: kininogen, HMW I #status predicted <MAT>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>


  Query Match             69.6%;  Score 112;  DB 2;  Length 639;
  Best Local Similarity   74.2%;  Pred. No. 3.6e-07;
  Matches   23;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           | |:| :|||||| |||||||| ||||  |||
Db     282 LGHSIAQLNAENNHTFYFKIDTVKKATSQVV 312


RESULT   9
KGRTM
major acute phase alpha-1 protein precursor - rat (fragment)
N;Contains: bradykinin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 12-Apr-1996
C;Accession: A01285
R;Cole, T.; Inglis, A.S.; Roxburgh, C.M.; Howlett, G.J.; Schreiber, G.
FEBS Lett. 182, 57-61, 1985
A;Title: Major acute phase alphal-protein of the rat is homologous to bovine kininoge
A;Reference number: A01285; MUID:85127561
A;Accession: A01285
A;Molecule type: mRNA
A;Residues: 1-423 <COL>
C;Comment: This plasma glycoprotein inhibits cysteine proteinases. During acute infla
r.
C;Superfamily: kininogen; cystatin homology
```

C;Keywords: bradykinin; cysteine proteinase inhibitor; duplication; glycoprotein; inflam
F;1-11/Domain: signal sequence (fragment) #status predicted <SIG>
F;12-423/Product: major acute phase alpha-1 protein #status predicted <MAT>
F;12-123/Domain: cystatin homology <CY1>
F;134-245/Domain: cystatin homology <CY2>
F;256-367/Domain: cystatin homology <CY3>
F;371-379/Product: bradykinin #status predicted <BDY>
F;12/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;161,197/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 61.5%; Score 99; DB 1; Length 423;
Best Local Similarity 64.5%; Pred. No. 1.3e-05;
Matches 20; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```
Qy       2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           | |:| :|||::|  |||||| |||| |||
Db     274 LGHSIARLNAQHNHIFYFKIDTVKKATSQVV 304
```

RESULT 10
KGRTT1
T-kininogen I precursor - rat
N;Alternate names: 73K protein; LMW kininogen T-I
N;Contains: bradykinin; T-kinin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 22-Jun-1999
C;Accession: A01286; D25486; A28526; PL0193; JQ0027; B25488; A28525; S68036
R;Furuto-Kato, S.; Matsumoto, A.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 12054-12059, 1985
A;Title: Primary structures of the mRNAs encoding the rat precursors for bradykinin and
nhibitor.
A;Reference number: A92496; MUID:86008264
A;Accession: A01286
A;Molecule type: mRNA
A;Residues: 1-430 <FUR>
A;Cross-references: GB:M11883; NID:g205084; PIDN:AAA41489.1; PID:g205085
R;Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A;Title: Differing expression patterns and evolution of the rat kininogen gene family.
A;Reference number: A92625; MUID:87137443
A;Accession: D25486
A;Molecule type: DNA
A;Residues: 375-430 <KIT>
R;Enjyoji, K.; Kato, H.; Hayashi, I.; Oh-ishi, S.; Iwanaga, S.
J. Biol. Chem. 263, 973-979, 1988
A;Title: Purification and characterization of rat T-kininogens isolated from plasma of a
A;Reference number: A92729; MUID:88087226
A;Accession: A28526
A;Molecule type: protein
A;Residues: 'E',20-48;376-430 <ENJ>
R;Kanda, S.; Sugiyama, K.; Takahashi, M.; Shumiya, S.; Tomino, S.; Nagase, S.
Jpn. J. Cancer Res. 81, 63-68, 1990
A;Title: Identification of a protein increasing in serum of Nagase analbuminemic rats be
A;Reference number: PL0193; MUID:90216390
A;Accession: PL0193
A;Molecule type: mRNA
A;Residues: 330-420,'R',422-429,'P' <KAN>
R;Anderson, K.P.; Croyle, M.L.; Lingrel, J.B.
Gene 81, 119-128, 1989
A;Title: Primary structure of a gene encoding rat T-kininogen.
A;Reference number: JQ0027; MUID:90034172
A;Accession: JQ0027
A;Molecule type: DNA
A;Residues: 1-60,'E',62-113,'R',115-165,'F',167-178,'TKI',182-211,'F',213-256,'S',258-38
A;Experimental source: strain Sprague-Dawley
R;Kageyama, R.; Kitamura, N.; Ohkubo, H.; Nakanishi, S.
J. Biol. Chem. 262, 2345-2351, 1987
A;Title: Differing utilization of homologous transcription initiation sites of rat K and
A;Reference number: A25488; MUID:87137465
A;Accession: B25488
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-48 <KAG>
A;Cross-references: GB:M14356; NID:g205090; PIDN:AAA41492.1; PID:g205091
R;Enjyoji, K.; Kato, H.; Hayashi, I.; Oh-ishi, S.; Iwanaga, S.
J. Biol. Chem. 263, 965-972, 1988
A;Title: Purification and characterization of two kinds of low molecular weight kinin
A;Reference number: A28525; MUID:88087225
A;Accession: A28525
A;Molecule type: protein
A;Residues: 376-430 <EN2>
R;Sierra, F.; Walter, R.; Vautravers, P.; Guigoz, Y.
Arch. Biochem. Biophys. 322, 333-338, 1995
A;Title: Identification of several isoforms of T-kininogen expressed in the liver of
A;Reference number: S68034; MUID:96032652
A;Accession: S68036
A;Molecule type: mRNA
A;Residues: 340-430 <SIE>
A;Experimental source: clone pSG17
C;Comment: At least three types of LMW kininogen precursors are present in rat plasma
ceding bradykinin.
C;Comment: T-kininogens contain T-kinin (I-S-bradykinin), a novel kinin isolated afte
d of an Arg or Lys, it is probably not released from its precursor by either tissue o
C;Comment: The T-kininogens are produced in response to an inflammatory stimulant.
C;Genetics:
A;Introns: 65/3; 102/3; 130/1; 187/3; 223/2; 252/1; 309/3; 345/3; 374/3; 398/3
C;Superfamily: kininogen; cystatin homology
C;Keywords: acute phase; bradykinin; cysteine proteinase inhibitor; duplication; glyc
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-430/Product: T-kininogen I #status experimental <MAT>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;378-386/Product: bradykinin #status predicted <BDY>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
F;82,126,168,204,326/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;83-94,107-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bonds: #sta

Query Match 60.9%; Score 98; DB 1; Length 430;
Best Local Similarity 64.5%; Pred. No. 1.8e-05;
Matches 20; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```
Qy       2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           | |:| :|||::|  |||||| |||| |||
Db     281 LGHSIAQLNAQHNHIFYFKIDTVKKATSQVV 311
```

RESULT 11
A23897
major acute phase alpha-1 protein (version 2) - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 20-Aug-1999
C;Accession: A23897; B23897
R;Anderson, K.P.; Heath, E.C.
J. Biol. Chem. 260, 12065-12071, 1985
A;Title: The relationship between rat major acute phase protein and the kininogens.
A;Reference number: A23897; MUID:86008266
A;Accession: A23897
A;Molecule type: protein
A;Residues: 1-14 <AND1>
A;Accession: B23897
A;Molecule type: mRNA
A;Residues: 5-430 <AND2>
A;Cross-references: GB:M11661; NID:g205307; PIDN:AAA41570.1; PID:g205308
A;Note: the authors translated the codon CTC for residue 410 as Arg, CTA for residue
C;Superfamily: kininogen; cystatin homology
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>

Query Match 60.9%; Score 98; DB 2; Length 430;
Best Local Similarity 64.5%; Pred. No. 1.8e-05;
Matches 20; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```
Qy       2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           | |:| :|||::|  ||||||  ||||  |||
Db     281 LGHSIAQLNAQHNHLFYFKIDTVKKATSQVV 311
```

RESULT 12
B28055
T-kininogen, LMW II precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 12-Dec-1997
C;Accession: B28055; E25486; B28526; C28526
R;Furuto-Kato, S.; Matsumoto, A.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 12054-12059, 1985
A;Title: Primary structures of the mRNAs encoding the rat precursors for bradykinin and
nhibitor.
A;Reference number: A92496; MUID:86008264
A;Accession: B28055
A;Molecule type: mRNA
A;Residues: 1-430 <FUR>
R;Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A;Title: Differing expression patterns and evolution of the rat kininogen gene family.
A;Reference number: A92625; MUID:87137443
A;Accession: E25486
A;Molecule type: DNA
A;Residues: 375-430 <KIT>
R;Enjyoji, K.; Kato, H.; Hayashi, I.; Oh-ishi, S.; Iwanaga, S.
J. Biol. Chem. 263, 973-979, 1988
A;Title: Purification and characterization of rat T-kininogens isolated from plasma of a
A;Reference number: A92729; MUID:88087226
A;Accession: B28526
A;Molecule type: protein
A;Residues: 'E',20-25,'MD',28-48,376-430 <ENJ>
A;Accession: C28526
A;Molecule type: protein
A;Residues: 'E',20-48;376-388,'R',390-419,'ER',422-430 <EN2>
C;Superfamily: kininogen; cystatin homology
C;Keywords: glycoprotein; pyroglutamic acid
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-430/Product: T-kininogen, LMW II #status experimental <MAT>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;82,126,168,204,326/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;83-94,107-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bonds: #status

```
  Query Match             60.9%;  Score 98;  DB 2;  Length 430;
  Best Local Similarity   64.5%;  Pred. No. 1.8e-05;
  Matches   20;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           | |:| :|||::|  ||||||  ||||  |||
Db     281 LGHSIAQLNAQHNHLFYFKIDTVKKATSQVV 311
```

RESULT 13
S49776
hypothetical protein YDR179w-a - yeast (Saccharomyces cerevisiae) (fragment)
N;Alternate names: hypothetical protein YD9395.13
C;Species: Saccharomyces cerevisiae
C;Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 29-Oct-1999
C;Accession: S49776; S51167
R;Murphy, L.; Harris, D.E.
submitted to the EMBL Data Library, November 1994
A;Reference number: S49764
A;Accession: S49776
A;Molecule type: DNA
A;Residues: 1-498 <MUR>
A;Cross-references: EMBL:Z46727; NID:g1289283; PIDN:CAA86685.1; PID:e223726; PID:g128929
A;Note: the sequence extends from the previous stop codon and does not begin with a star
A;Accession: S51167
A;Molecule type: DNA
A;Residues: 231-498 <MUW>
A;Cross-references: EMBL:Z46727; NID:g1289283; PID:e223644; PID:g1289296; GSPDB:GN000
A;Note: the reading frame starts at the first ATG codon
C;Genetics:
A;Gene: MIPS:YDR179w-a
A;Map position: 4R

```
  Query Match             32.6%;  Score 52.5;  DB 2;  Length 498;
  Best Local Similarity   41.4%;  Pred. No. 25;
  Matches   12;  Conservative    5;  Mismatches   11;  Indels    1;  Gaps    1;

Qy       1 TLTHTITKLNAENNATFYFKIDNVKKARV 29
           || | |:|||| |    |   ||: : ::
Db     374 TLQHWISKLNALNR-VMYRTFDNIVQTKI 401
```

RESULT 14
B97948
hypothetical protein ABC-NBD-truncation [imported] - Streptococcus pneumoniae (strain
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C;Accession: B97948
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczak, L.; Burgett, S.; DeHoff, B.S.
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: B97948
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-217 <KUR>
A;Cross-references: GB:AE007317; PIDN:AAK99414.1; PID:g15458193; GSPDB:GN00174
C;Genetics:
A;Gene: ABC-NBD-truncation

```
  Query Match             32.3%;  Score 52;  DB 2;  Length 217;
  Best Local Similarity   45.5%;  Pred. No. 13;
  Matches   10;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       8 KLNAENNATFYFKIDNVKKARV 29
           |::  ||  || |  |: |:||
Db      28 KMSLYNNLLFYSKFYNISKSRV 49
```

RESULT 15
A44943
heat shock protein 83 - Leishmania mexicana amazonensis
C;Species: Leishmania mexicana amazonensis
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Aug-1999
C;Accession: A44943
R;Shapira, M.; Pedraza, G.
Mol. Biochem. Parasitol. 42, 247-256, 1990
A;Title: Sequence analysis and transcriptional activation of heat shock protein 83 of
A;Reference number: A44943; MUID:91101661
A;Accession: A44943
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-701 <SHA>
A;Cross-references: GB:M92926; GB:M34292; NID:g159351; PIDN:AAA29250.1; PID:g159354
C;Superfamily: heat shock protein 90

```
  Query Match             32.3%;  Score 52;  DB 2;  Length 701;
  Best Local Similarity   38.5%;  Pred. No. 42;
  Matches   10;  Conservative    6;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAENNATFYFKIDNVKK 26
```

```
          ||   :|:: || |: :|   |: ||
Db    448 TLKDYVTRMKAEQNSIYYITGDSKKK 473
```

Search completed: July 1, 2002, 16:20:39
Job time: 203 sec

GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:30:13 ; Search time 21.51 Seconds
(without alignments)
57.602 Million cell updates/sec

Title: US-09-461-061A-4
Perfect score: 161
Sequence: 1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 161 | 100.0 | 644 | 1 | KNG_HUMAN | P01042 homo sapien |
| 2 | 116 | 72.0 | 434 | 1 | KNL2_BOVIN | P01047 bos taurus |
| 3 | 116 | 72.0 | 619 | 1 | KNH2_BOVIN | P01045 bos taurus |
| 4 | 112 | 69.6 | 436 | 1 | KNL1_BOVIN | P01046 bos taurus |
| 5 | 112 | 69.6 | 621 | 1 | KNH1_BOVIN | P01044 bos taurus |
| 6 | 112 | 69.6 | 639 | 1 | KNG_RAT | P08934 rattus norv |
| 7 | 98 | 60.9 | 430 | 1 | KNT1_RAT | P01048 rattus norv |
| 8 | 98 | 60.9 | 430 | 1 | KNT2_RAT | P08932 rattus norv |
| 9 | 98 | 60.9 | 661 | 1 | KNG_MOUSE | O08677 mus musculu |
| 10 | 61 | 37.9 | 144 | 1 | CYTF_MOUSE | O89098 mus musculu |
| 11 | 52 | 32.3 | 316 | 1 | UFD1_DROME | Q9vtf9 drosophila |
| 12 | 52 | 32.3 | 701 | 1 | HS83_LEIAM | P27741 leishmania |
| 13 | 50 | 31.1 | 306 | 1 | YDS4_SCHPO | O14180 schizosacch |
| 14 | 49.5 | 30.7 | 352 | 1 | A2HS_RAT | P24090 rattus norv |
| 15 | 49 | 30.4 | 358 | 1 | T2H2_HAEPA | P36433 haemophilus |
| 16 | 48 | 29.8 | 145 | 1 | CYTF_HUMAN | O76096 homo sapien |
| 17 | 47.5 | 29.5 | 392 | 1 | EFT3_STRCO | P40175 streptomyce |
| 18 | 47.5 | 29.5 | 452 | 1 | EFTU_BOVIN | P49410 bos taurus |
| 19 | 47.5 | 29.5 | 452 | 1 | EFTU_HUMAN | P49411 homo sapien |
| 20 | 47.5 | 29.5 | 645 | 1 | T3MO_ECOLI | P12364 escherichia |
| 21 | 47 | 29.2 | 111 | 1 | CYT_BITAR | P08935 bitis ariet |
| 22 | 47 | 29.2 | 224 | 1 | KCY_UREPA | Q9pqe9 ureaplasma |
| 23 | 47 | 29.2 | 348 | 1 | LLY_LEGPN | Q53407 legionella |
| 24 | 46.5 | 28.9 | 345 | 1 | A2HS_MOUSE | P29699 mus musculu |
| 25 | 46 | 28.6 | 337 | 1 | TH23_TRYBB | Q09039 trypanosoma |
| 26 | 46 | 28.6 | 400 | 1 | YIEL_ECOLI | P31471 escherichia |
| 27 | 46 | 28.6 | 419 | 1 | EXON_NPVAC | P24731 autographa |
| 28 | 46 | 28.6 | 527 | 1 | TH11_TRYBB | Q06221 trypanosoma |
| 29 | 46 | 28.6 | 615 | 1 | YKZ9_YEAST | P36115 saccharomyc |
| 30 | 46 | 28.6 | 620 | 1 | AFR1_YEAST | P33304 saccharomyc |
| 31 | 46 | 28.6 | 624 | 1 | HTPG_CLOAB | Q97e05 clostridium |
| 32 | 46 | 28.6 | 1220 | 1 | SLN1_YEAST | P39928 saccharomyc |
| 33 | 45.5 | 28.3 | 668 | 1 | VNCS_FPV19 | P24842 feline panl |
| 34 | 45.5 | 28.3 | 668 | 1 | VNCS_MEVA | P27438 mink enteri |
| 35 | 45.5 | 28.3 | 668 | 1 | VNCS_PAVCN | P12929 canine parv |
| 36 | 45 | 28.0 | 57 | 1 | CGKB_ALTCA | P43479 alteromonas |
| 37 | 45 | 28.0 | 108 | 1 | KV6K_MOUSE | P04945 mus musculu |
| 38 | 45 | 28.0 | 148 | 1 | CYTC_BOVIN | P01035 bos taurus |
| 39 | 45 | 28.0 | 149 | 1 | CYTM_HUMAN | Q15828 homo sapien |
| 40 | 45 | 28.0 | 264 | 1 | YE54_SCHPO | O14171 schizosacch |
| 41 | 45 | 28.0 | 312 | 1 | TRUB_BUCAI | P57456 buchnera ap |
| 42 | 45 | 28.0 | 389 | 1 | EFT3_STRRA | P29544 streptomyce |
| 43 | 45 | 28.0 | 394 | 1 | EFTU_RECAM | O21245 reclinomona |
| 44 | 45 | 28.0 | 446 | 1 | GLNA_METMP | O59648 methanococc |
| 45 | 45 | 28.0 | 850 | 1 | D7_DICDI | P54682 dictyosteli |

ALIGNMENTS

RESULT 1
KNG_HUMAN
ID KNG_HUMAN STANDARD; PRT; 644 AA.
AC P01042; P01043;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Kininogen precursor (Alpha-2-thiol proteinase inhibitor) [Contains:
DE Bradykinin].
GN KNG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC TISSUE=Liver;
RX MEDLINE=85234582; PubMed=2989293;
RA Takagaki Y., Kitamura N., Nakanishi S.;
RT "Cloning and sequence analysis of cDNAs for human high molecular
RT weight and low molecular weight prekininogens. Primary structures of
RT two human prekininogens.";
RL J. Biol. Chem. 260:8601-8609(1985).
RN [2]
RP GENE STRUCTURE.
RX MEDLINE=85234583; PubMed=2989294;
RA Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T.,
RA Nakanishi S.;
RT "Structural organization of the human kininogen gene and a model for
RT its evolution.";
RL J. Biol. Chem. 260:8610-8617(1985).
RN [3]
RP SEQUENCE OF 1-401 FROM N.A.
RX MEDLINE=85122621; PubMed=6441591;
RA Ohkubo I., Kurachi K., Takasawa T., Shiokawa H., Sasaki M.;
RT "Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and
RT its identity with low molecular weight kininogen.";
RL Biochemistry 23:5691-5697(1984).
RN [4]
RP SEQUENCE OF 379-644.
RX MEDLINE=86030270; PubMed=4054110;
RA Lottspeich F., Kellermann J., Henschen A., Foertsch B.,
RA Mueller-Esterl W.;
RT "The amino acid sequence of the light chain of human high-molecular-
RT mass kininogen.";
RL Eur. J. Biochem. 152:307-314(1985).
RN [5]
RP SEQUENCE OF 381-389.
RX MEDLINE=90255622; PubMed=4952632;
RA Pierce J.V.;
RT "Structural features of plasma kinins and kininogens.";
RL Fed. Proc. 27:52-57(1968).
RN [6]
RP DISULFIDE BONDS.
RA Sueyoshi T., Miyata T., Kato H., Iwanaga S.;
RT "Disulfide bonds in bovine HMW kininogens.";

```
RL   Seikagaku 56:808-808(1984).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC       HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
CC       FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-
CC       INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
CC       BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
CC       PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
CC       CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
CC       DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
CC       MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
CC       PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
CC       OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
CC       A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
CC       INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
CC       KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
CC       CLOTTING.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; K02566; AAA35497.1; -.
DR   EMBL; M11437; AAB59550.1; -.
DR   EMBL; M11438; AAB59550.1; JOINED.
DR   EMBL; M11521; AAB59550.1; JOINED.
DR   EMBL; M11522; AAB59550.1; JOINED.
DR   EMBL; M11523; AAB59550.1; JOINED.
DR   EMBL; M11524; AAB59550.1; JOINED.
DR   EMBL; M11525; AAB59550.1; JOINED.
DR   EMBL; M11526; AAB59550.1; JOINED.
DR   EMBL; M11527; AAB59550.1; JOINED.
DR   EMBL; M11528; AAB59550.1; JOINED.
DR   EMBL; M11437; AAB59551.1; -.
DR   EMBL; M11438; AAB59551.1; JOINED.
DR   EMBL; M11521; AAB59551.1; JOINED.
DR   EMBL; M11522; AAB59551.1; JOINED.
DR   EMBL; M11523; AAB59551.1; JOINED.
DR   EMBL; M11524; AAB59551.1; JOINED.
DR   EMBL; M11525; AAB59551.1; JOINED.
DR   EMBL; M11526; AAB59551.1; JOINED.
DR   EMBL; M11527; AAB59551.1; JOINED.
DR   EMBL; M11528; AAB59551.1; JOINED.
DR   PIR; A01279; KGHUH1.
DR   PIR; A25276; A25276.
DR   PIR; A01280; KGHUL1.
DR   PIR; B25276; B25276.
DR   PIR; S02482; S02482.
DR   SWISS-2DPAGE; P01042; HUMAN.
DR   MIM; 228960; -.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW   Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW   Alternative splicing.
FT   SIGNAL        1     18
FT   CHAIN         19    644       KININOGEN.
FT   CHAIN         19    380       KININOGEN HEAVY CHAIN.
FT   PEPTIDE      381    389       BRADYKININ.
FT   CHAIN        390    644       KININOGEN LIGHT CHAIN.
FT   DOMAIN        19    136       CYSTATIN-LIKE 1.
FT   DOMAIN       137    258       CYSTATIN-LIKE 2.
FT   DOMAIN       259    380       CYSTATIN-LIKE 3.
FT   DOMAIN       420    510       HIS-RICH (ASSOCIATED WITH CLOTTING
FT                                 ACTIVITY).
FT   REPEAT       420    449
FT   REPEAT       450    479
FT   REPEAT       480    510
FT   MOD_RES       19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   DISULFID      28    614       INTERCHAIN.
FT   DISULFID      83     94
FT   DISULFID     107    126
FT   DISULFID     142    145
FT   DISULFID     206    218
FT   DISULFID     229    248
FT   DISULFID     264    267
FT   DISULFID     328    340
FT   DISULFID     351    370
FT   CARBOHYD      48     48       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     169    169       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     205    205       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     294    294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     401    401       O-LINKED.
FT   CARBOHYD     533    533       O-LINKED.
FT   CARBOHYD     542    542       O-LINKED.
FT   CARBOHYD     546    546       O-LINKED.
FT   CARBOHYD     557    557       O-LINKED.
FT   CARBOHYD     571    571       O-LINKED.
FT   CARBOHYD     577    577       O-LINKED.
FT   CARBOHYD     593    593       O-LINKED.
FT   CARBOHYD     628    628       O-LINKED.
FT   VARSPLIC     402    427       VSPPHTSMAPAQDEERDSGKEQGHTR -> SHLRSCEYKGR
FT                                 PPKAGAEPASEREVS (IN ISOFORM LMW).
FT   VARSPLIC     428    644       MISSING (IN ISOFORM LMW).
FT   CONFLICT     593    593       T -> I (IN REF. 1).
SQ   SEQUENCE   644 AA;  71945 MW;  3132B4CBAF8FBB7E CRC64;


  Query Match            100.0%;  Score 161;  DB 1;  Length 644;
  Best Local Similarity  100.0%;  Pred. No. 4.5e-15;
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           ||||||||||||||||||||||||||||||||
Db     281 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 312


RESULT   2
KNL2_BOVIN
ID   KNL2_BOVIN     STANDARD;      PRT;   434 AA.
AC   P01047;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Kininogen, LMW II precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83117859; PubMed=6572010;
RA   Nawa H., Kitamura N., Hirose T., Asai M., Inayama S., Nakanishi S.;
RT   "Primary structures of bovine liver low molecular weight kininogen
RT   precursors and their two mRNAs.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:90-94(1983).
RN   [2]
```

```
RP   SEQUENCE OF 19-376.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (3) THE
CC       ACTIVE PEPTIDE KALLIDIN THAT IS RELEASED FROM LMW-KININOGEN SHOWS
CC       A VARIETY OF PHYSIOLOGICAL EFFECTS: (3A) INFLUENCE IN SMOOTH
CC       MUSCLE CONTRACTION, (3B) INDUCTION OF HYPOTENSION, (3C)
CC       NATRIURESIS AND DIURESIS (KIDNEY).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS: HMW II AND LMW II KININOGEN PRECURSORS ARE
CC       PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
CC       SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
CC       TO RESIDUE 398.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- MISCELLANEOUS: LMW-KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT
CC       INVOLVED IN BLOOD CLOTTING.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00427; CAA23710.1; -.
DR   PIR; A01284; KGBOL2.
DR   HSSP; P01038; 1A90.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   Pfam; PF00031; cystatin; 3.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Signal.
FT   SIGNAL        1     18
FT   CHAIN        19    434       KININOGEN, LMW II.
FT   CHAIN        19    376       HEAVY CHAIN.
FT   PEPTIDE     378    386       BRADYKININ.
FT   CHAIN       387    434       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    256       CYSTATIN-LIKE 2.
FT   DOMAIN      257    376       CYSTATIN-LIKE 3.
FT   MOD_RES      19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    136    136       O-LINKED (PARTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .); OR 169.
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    280    280       N-LINKED (GLCNAC. . .).
FT   DISULFID     27    404       INTERCHAIN.
FT   DISULFID     82     93
FT   DISULFID    106    125
FT   DISULFID    141    144
FT   DISULFID    205    217
FT   DISULFID    228    247
FT   DISULFID    261    264
FT   DISULFID    325    337
FT   DISULFID    348    367
SQ   SEQUENCE   434 AA;  48148 MW;  73A7079DE3E03430 CRC64;

 Query Match           72.0%;  Score 116;  DB 1;  Length 434;
 Best Local Similarity 74.2%;  Pred. No. 6e-09;
   Matches   23;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           | |:| ||||||:: ||||||| |||| ||||
Db     279 LNHSIAKLNAEHDGTFYFKIDTVKKATVQVV 309


RESULT   3
KNH2_BOVIN
ID   KNH2_BOVIN     STANDARD;      PRT;   619 AA.
AC   P01045;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Kininogen, HMW II precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84014106; PubMed=6571699;
RA   Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
RT   "A single gene for bovine high molecular weight and low molecular
RT   weight kininogens.";
RL   Nature 305:545-549(1983).
RN   [2]
RP   SEQUENCE OF 19-376.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
RN   [3]
RP   SEQUENCE OF 376-391.
RX   MEDLINE=70180420; PubMed=4986212;
RA   Kato H., Nagasawa S., Suzuki T.;
RT   "Studies on the structure of bovine kininogen: cleavages of disulfide
RT   bonds and of methionyl bonds in kininogen-II.";
RL   J. Biochem. 67:313-323(1970).
RN   [4]
RP   SEQUENCE OF 387-455.
RX   MEDLINE=76260155; PubMed=956151;
RA   Han Y.N., Kato H., Iwanaga S., Suzuki T.;
RT   "Primary structure of bovine plasma high-molecular-weight kininogen.
RT   The amino acid sequence of a glycopeptide portion (fragment 1)
RT   following the C-terminus ot the bradykinin moiety.";
RL   J. Biochem. 79:1201-1222(1976).
RN   [5]
RP   SEQUENCE OF 456-496.
RX   MEDLINE=75170265; PubMed=1169237;
RA   Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
RT   "Studies on the primary structure of bovine high-molecular-weight
RT   kininogen. Amino acid sequence of a fragment ('histidine-rich
RT   peptide') released by plasma kallikrein.";
RL   J. Biochem. 77:55-68(1975).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC       HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
CC       TO FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
CC       PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE
CC       PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
CC       A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
CC       MUSCLE CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C)
CC       NATRIURESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
CC       (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
CC       IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
CC       RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
CC       (4F) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
```

```
CC       ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
CC       ACTION).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS: HMW II AND LMW II KININOGEN PRECURSORS ARE
CC       PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
CC       SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
CC       TO RESIDUE 398.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V01492; CAA24736.1; -.
DR   EMBL; V01492; CAA24737.1; ALT_SEQ.
DR   PIR; A01282; KGBOH2.
DR   PIR; B29559; B29559.
DR   HSSP; P01038; 1A90.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Blood coagulation; Signal;
KW   Inflammatory response.
FT   SIGNAL        1     18
FT   CHAIN        19    619       KININOGEN, HMW II.
FT   CHAIN        19    376       HEAVY CHAIN.
FT   PEPTIDE     378    386       BRADYKININ.
FT   CHAIN       387    619       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    256       CYSTATIN-LIKE 2.
FT   DOMAIN      257    376       CYSTATIN-LIKE 3.
FT   MOD_RES      19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    136    136       O-LINKED (PARTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .); OR 169.
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    280    280       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    400    400       O-LINKED.
FT   DISULFID     27    589       INTERCHAIN.
FT   DISULFID     82     93
FT   DISULFID    106    125
FT   DISULFID    141    144
FT   DISULFID    205    217
FT   DISULFID    228    247
FT   DISULFID    261    264
FT   DISULFID    325    337
FT   DISULFID    348    367
FT   VARIANT     398    398       T -> P.
FT   VARIANT     401    401       L -> V.
FT   VARIANT     454    454       H -> K.
SQ   SEQUENCE   619 AA;  68710 MW;  F04320A8EB0EE0DA CRC64;


  Query Match          72.0%;  Score 116;  DB 1;  Length 619;
  Best Local Similarity  74.2%;  Pred. No. 8.7e-09;
  Matches   23;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           | |:| ||||::| ||||||| |||| ||||
Db     279 LNHSIAKLNAEHDGTFYFKIDTVKKATVQVV 309
```

```
RESULT   4
KNL1_BOVIN
ID   KNL1_BOVIN     STANDARD;      PRT;   436 AA.
AC   P01046;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Kininogen, LMW I precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83117859; PubMed=6572010;
RA   Nawa H., Kitamura N., Hirose T., Asai M., Inayama S., Nakanishi S.;
RT   "Primary structures of bovine liver low molecular weight kininogen
RT   precursors and their two mRNAs.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:90-94(1983).
RN   [2]
RP   SEQUENCE OF 19-378.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (3) THE
CC       ACTIVE PEPTIDE KALLIDIN THAT IS RELEASED FROM LMW-KININOGEN SHOWS
CC       A VARIETY OF PHYSIOLOGICAL EFFECTS: (3A) INFLUENCE IN SMOOTH
CC       MUSCLE CONTRACTION, (3B) INDUCTION OF HYPOTENSION, (3C)
CC       NATRIURESIS AND DIURESIS (KIDNEY).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS: HMW I AND LMW I KININOGEN PRECURSORS ARE
CC       PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
CC       SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
CC       TO RESIDUE 400.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- MISCELLANEOUS: LMW-KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT
CC       INVOLVED IN BLOOD CLOTTING.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00426; CAA23709.1; -.
DR   EMBL; J00010; AAA30604.1; -.
DR   PIR; A01283; KGBOL1.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   Pfam; PF00031; cystatin; 3.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Signal.
FT   SIGNAL        1     18
FT   CHAIN        19    436       KININOGEN, LMW I.
FT   CHAIN        19    378       HEAVY CHAIN.
FT   PEPTIDE     380    388       BRADYKININ.
FT   CHAIN       389    436       LIGHT CHAIN.
```

```
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    257       CYSTATIN-LIKE 2.
FT   DOMAIN      258    378       CYSTATIN-LIKE 3.
FT   MOD_RES      19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    136    136       O-LINKED (PARTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .); OR 169.
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .).
FT   DISULFID     27    406       INTERCHAIN.
FT   DISULFID     82     93
FT   DISULFID    106    125
FT   DISULFID    141    144
FT   DISULFID    205    217
FT   DISULFID    228    247
FT   DISULFID    263    266
FT   DISULFID    327    339
FT   DISULFID    350    369
FT   CONFLICT    295    295       A -> T (IN REF. 1; CAA23709).
SQ   SEQUENCE   436 AA;  48427 MW;  F01F7EB6814BCE6C CRC64;


  Query Match             69.6%;  Score 112;  DB 1;  Length 436;
  Best Local Similarity   71.0%;  Pred. No. 2.2e-08;
  Matches   22;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           |:|:| |||||::  |||||| |||| ||||
Db     281 LSHSIAKLNAEHDGAFYFKIDTVKKATVQVV 311


RESULT   5
KNH1_BOVIN
ID   KNH1_BOVIN     STANDARD;      PRT;   621 AA.
AC   P01044;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Kininogen, HMW I precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84014106; PubMed=6571699;
RA   Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
RT   "A single gene for bovine high molecular weight and low molecular
RT   weight kininogens.";
RL   Nature 305:545-549(1983).
RN   [2]
RP   SEQUENCE OF 19-378.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
RN   [3]
RP   SEQUENCE OF 378-393.
RX   MEDLINE=70180420; PubMed=4986212;
RA   Kato H., Nagasawa S., Suzuki T.;
RT   "Studies on the structure of bovine kininogen: cleavages of disulfide
RT   bonds and of methionyl bonds in kininogen-II.";
RL   J. Biochem. 67:313-323(1970).
RN   [4]
RP   SEQUENCE OF 458-498.
RX   MEDLINE=75170265; PubMed=1169237;
RA   Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
RT   "Studies on the primary structure of bovine high-molecular-weight
RT   kininogen. Amino acid sequence of a fragment ('histidine-rich
RT   peptide') released by plasma kallikrein.";
RL   J. Biochem. 77:55-68(1975).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC       HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
CC       TO FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
CC       PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE
CC       PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
CC       A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
CC       MUSCLE CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C)
CC       NATRIURESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
CC       (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
CC       IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
CC       RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
CC       (4F) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
CC       ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
CC       ACTION).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS: HMW I AND LMW I KININOGEN PRECURSORS ARE
CC       PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
CC       SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
CC       TO RESIDUE 400.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; V01491; CAA24735.1; -.
DR   PIR; A01281; KGBOH1.
DR   PIR; A29559; A29559.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Blood coagulation;
KW   Inflammatory response; Signal.
FT   SIGNAL        1     18       PROBABLE.
FT   CHAIN        19    621       KININOGEN, HMW I.
FT   CHAIN        19    378       HEAVY CHAIN.
FT   PEPTIDE     380    388       BRADYKININ.
FT   CHAIN       389    621       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    257       CYSTATIN-LIKE 2.
FT   DOMAIN      258    378       CYSTATIN-LIKE 3.
FT   MOD_RES      19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    136    136       O-LINKED (PARTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .); OR 169.
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .).
FT   DISULFID     27    591       INTERCHAIN.
FT   DISULFID     82     93
FT   DISULFID    106    125
FT   DISULFID    141    144
FT   DISULFID    205    217
FT   DISULFID    228    247
FT   DISULFID    263    266
FT   DISULFID    327    339
FT   DISULFID    350    369
SQ   SEQUENCE   621 AA;  68890 MW;  D16850BEFE3C55CD CRC64;
```

```
  Query Match             69.6%;  Score 112;  DB 1;  Length 621;
  Best Local Similarity   71.0%;  Pred. No. 3.2e-08;
  Matches   22;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           |:|:| |||||::  |||||| |||| ||||
Db     281 LSHSIAKLNAEHDGAFYFKIDTVKKATVQVV 311


RESULT   6
KNG_RAT
ID   KNG_RAT        STANDARD;      PRT;   639 AA.
AC   P08934; P08933;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Kininogen precursor [Contains: Bradykinin].
GN   KNG.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RX   MEDLINE=87137443; PubMed=3029068;
RA   Kitagawa H., Kitamura N., Hayashida H., Miyata T., Nakanishi S.;
RT   "Differing expression patterns and evolution of the rat kininogen
RT   gene family.";
RL   J. Biol. Chem. 262:2190-2198(1987).
RN   [2]
RP   SEQUENCE FROM N.A. (LMW ISOFORM).
RX   MEDLINE=86008264; PubMed=2413018;
RA   Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT   "Primary structures of the mRNAs encoding the rat precursors for
RT   bradykinin and T-kinin. Structural relationship of kininogens with
RT   major acute phase protein and alpha 1-cysteine proteinase
RT   inhibitor.";
RL   J. Biol. Chem. 260:12054-12059(1985).
RN   [3]
RP   SEQUENCE OF 1-65 FROM N.A.
RC   STRAIN=BUFFALO;
RX   MEDLINE=87250580; PubMed=2439509;
RA   Fung W.-P., Schreiber G.;
RT   "Structure and expression of the genes for major acute phase alpha 1-
RT   protein (thiostatin) and kininogen in the rat.";
RL   J. Biol. Chem. 262:9298-9308(1987).
RN   [4]
RP   SEQUENCE OF 1-41 FROM N.A.
RC   STRAIN=WISTAR; TISSUE=Liver;
RX   MEDLINE=87137465; PubMed=3818598;
RA   Kageyama R., Kitamura N., Ohkubo H., Nakanishi S.;
RT   "Differing utilization of homologous transcription initiation sites
RT   of rat K and T kininogen genes under inflammation condition.";
RL   J. Biol. Chem. 262:2345-2351(1987).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC       HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
CC       FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-
CC       INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
CC       BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
CC       PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
CC       CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
CC       DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
CC       MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
CC       PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
CC       OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
CC       A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
CC       INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
CC       KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
CC       CLOTTING.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- MISCELLANEOUS: RAT EXPRESS FOUR TYPES OF KININOGENS: THE CLASSICAL
CC       HMW/LMW KININOGENS AND TWO ADDITIONAL LMW-LIKE KININOGENS: T-I AND
CC       T-II.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29428; AAA41486.1; -.
DR   EMBL; M11884; AAA41487.1; -.
DR   EMBL; M14369; AAA41484.1; -.
DR   EMBL; M14369; AAA41485.1; ALT_SEQ.
DR   EMBL; M16455; AAA41482.1; -.
DR   PIR; A25486; A25486.
DR   PIR; A28055; A28055.
DR   HSSP; P01040; 1DVD.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW   Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW   Alternative splicing; Multigene family.
FT   SIGNAL       1     18
FT   CHAIN       19    639       KININOGEN.
FT   CHAIN       19    380       KININOGEN HEAVY CHAIN.
FT   PEPTIDE    381    389       BRADYKININ.
FT   CHAIN      390    639       KININOGEN LIGHT CHAIN.
FT   DOMAIN      19    136       CYSTATIN-LIKE 1.
FT   DOMAIN     137    258       CYSTATIN-LIKE 2.
FT   DOMAIN     259    380       CYSTATIN-LIKE 3.
FT   DOMAIN     439    514       HIS-RICH.
FT   DISULFID    28    609       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    83     94       BY SIMILARITY.
FT   DISULFID   107    126       BY SIMILARITY.
FT   DISULFID   142    145       BY SIMILARITY.
FT   DISULFID   206    218       BY SIMILARITY.
FT   DISULFID   229    248       BY SIMILARITY.
FT   DISULFID   264    267       BY SIMILARITY.
FT   DISULFID   328    340       BY SIMILARITY.
FT   DISULFID   351    370       BY SIMILARITY.
FT   CARBOHYD    82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   127    127       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   169    169       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   205    205       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   294    294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   529    529       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   402    433       VSPSYIARVQEERDPGNEQGPIHGHGWLHAKQ -> RLLNS
FT                               CEYKGRLLKAGAGPAPERQAEASTVTP (IN ISOFORM
FT                               LMW).
FT   VARSPLIC   434    639       MISSING (IN ISOFORM LMW).
FT   CONFLICT    61     61       E -> K (IN REF. 2).
SQ   SEQUENCE   639 AA;  70933 MW;  D3172DF94FF56AF5 CRC64;


  Query Match             69.6%;  Score 112;  DB 1;  Length 639;
  Best Local Similarity   74.2%;  Pred. No. 3.3e-08;
  Matches   23;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;
```

```
Qy      12 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           | |:| :|||||| ||||||| |||| |||
Db     282 LGHSIAQLNAENNHTFYFKIDTVKKATSQVV 312


RESULT   7
KNT1_RAT
ID   KNT1_RAT       STANDARD;      PRT;   430 AA.
AC   P01048; P04081;
DT   01-NOV-1986 (Rel. 03, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   T-kininogen I precursor (Major acute phase protein) (Alpha-1-MAP)
DE   (Thiostatin) [Contains: T-kinin].
GN   MAP1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86008264; PubMed=2413018;
RA   Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT   "Primary structures of the mRNAs encoding the rat precursors for
RT   bradykinin and T-kinin. Structural relationship of kininogens with
RT   major acute phase protein and alpha 1-cysteine proteinase
RT   inhibitor.";
RL   J. Biol. Chem. 260:12054-12059(1985).
RN   [2]
RP   SEQUENCE OF 5-430 FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=86008266; PubMed=2413019;
RA   Anderson K.P., Heath E.C.;
RT   "The relationship between rat major acute phase protein and the
RT   kininogens.";
RL   J. Biol. Chem. 260:12065-12071(1985).
RN   [3]
RP   SEQUENCE OF 7-430 FROM N.A.
RX   MEDLINE=85127561; PubMed=2578992;
RA   Cole T., Inglis A.S., Roxburgh C.M., Howlett G.J., Schreiber G.;
RT   "Major acute phase alpha 1-protein of the rat is homologous to bovine
RT   kininogen and contains the sequence for bradykinin: its synthesis is
RT   regulated at the mRNA level.";
RL   FEBS Lett. 182:57-61(1985).
RN   [4]
RP   SEQUENCE OF 1-65 FROM N.A.
RX   MEDLINE=87250580; PubMed=2439509;
RA   Fung W.-P., Schreiber G.;
RT   "Structure and expression of the genes for major acute phase alpha 1-
RT   protein (thiostatin) and kininogen in the rat.";
RL   J. Biol. Chem. 262:9298-9308(1987).
CC   -!- FUNCTION: KININOGENS ARE PLASMA GLYCOPROTEINS WITH A NUMBER OF
CC       FUNCTIONS: (1) AS PRECURSOR OF THE ACTIVE PEPTIDE BRADYKININ THEY
CC       EFFECT SMOOTH MUSCLE CONTRACTION, INDUCTION OF HYPOTENSION AND
CC       INCREASE OF VASCULAR PERMEABILITY. (2) THEY PLAY A ROLE IN BLOOD
CC       COAGULATION BY HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND
CC       FACTOR XI NEXT TO FACTOR XII. (3) THEY ARE INHIBITOR OF THIOL
CC       PROTEASES.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- INDUCTION: IN RESPONSE TO AN INFLAMMATORY STIMULANT. T-KININOGEN
CC       II SYNTHESIS IS INDUCED AND THE PLASMA CONCENTRATION OF
CC       T-KININOGEN I IS RAISED.
CC   -!- PTM: AS T-KININ IS PRECEDED BY A MET INSTEAD OF AN ARG OR LYS, IT
CC       IS NOT RELEASED FROM ITS PRECURSOR BY EITHER TISSUE OR PLASMA
CC       KALLIKREIN.
CC   -!- MISCELLANEOUS: RAT EXPRESS FOUR TYPES OF KININOGENS: THE CLASSICAL
CC       HMW AND LMW KININOGENS PRODUCED BY ALTERNATIVE SPLICING OF THE
CC       SAME GENE, AND TWO ADDITIONAL LMW-LIKE KININOGENS: T-I AND T-II.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   -!- CAUTION: IN ADDITION TO THE CONFLICTS DESCRIBED IN THE FEATURE
CC       TABLE, REF.2. SEQUENCE DIFFERS FROM THAT SHOWN IN POSITIONS: 257,
CC       262,268,269,295,314,315,331,332, AND 389. IN ALL THOSE POSITIONS
CC       THE ALTERNATE AMINO-ACID IS THE ONE PRESENT IN T-II KININOGEN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M11883; AAA41489.1; -.
DR   EMBL; M11661; AAA41570.1; -.
DR   EMBL; M16454; AAA41568.1; -.
DR   EMBL; X02299; CAA26162.1; ALT_SEQ.
DR   PIR; A01286; KGRTT1.
DR   PIR; A01285; KGRTM.
DR   PIR; A23897; A23897.
DR   PIR; A27115; A27115.
DR   GlycoSuiteDB; P01048; -.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   Pfam; PF00031; cystatin; 3.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Multigene family;
KW   Thiol protease inhibitor; Bradykinin; Acute phase; Signal.
FT   SIGNAL        1     18
FT   CHAIN        19    430       KININOGEN, T-I.
FT   CHAIN        19    375       HEAVY CHAIN.
FT   PEPTIDE     376    386       T-KININ.
FT   CHAIN       387    430       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    257       CYSTATIN-LIKE 2.
FT   DOMAIN      258    375       CYSTATIN-LIKE 3.
FT   DISULFID     28    404       INTERCHAIN (BY SIMILARITY).
FT   DISULFID     83     94       BY SIMILARITY.
FT   DISULFID    107    125       BY SIMILARITY.
FT   DISULFID    141    144       BY SIMILARITY.
FT   DISULFID    205    217       BY SIMILARITY.
FT   DISULFID    228    247       BY SIMILARITY.
FT   DISULFID    263    266       BY SIMILARITY.
FT   DISULFID    327    339       BY SIMILARITY.
FT   DISULFID    350    369       BY SIMILARITY.
FT   CARBOHYD     82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    326    326       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     26     28       LNC -> MDR (IN REF. 2).
FT   CONFLICT     55     55       V -> L (IN REF. 2).
FT   CONFLICT     61     61       E -> K (IN REF. 1).
FT   CONFLICT     83     83       C -> Y (IN REF. 3).
FT   CONFLICT    166    166       S -> F (IN REF. 2 AND 3).
FT   CONFLICT    179    181       REV -> TKI (IN REF. 2).
FT   CONFLICT    193    193       N -> D (IN REF. 2).
FT   CONFLICT    212    212       S -> F (IN REF. 2).
FT   CONFLICT    214    214       R -> H (IN REF. 3).
FT   CONFLICT    229    229       T -> R (IN REF. 2).
FT   CONFLICT    233    233       H -> Y (IN REF. 2).
FT   CONFLICT    257    257       E -> S (IN REF. 2).
FT   CONFLICT    262    262       N -> K (IN REF. 2).
FT   CONFLICT    264    264       R -> F (IN REF. 2).
FT   CONFLICT    268    269       RE -> KN (IN REF. 2).
FT   CONFLICT    295    295       I -> L (IN REF. 2).
FT   CONFLICT    314    315       VI -> TK (IN REF. 2).
FT   CONFLICT    331    332       SK -> TN (IN REF. 2).
FT   CONFLICT    389    389       R -> Q (IN REF. 2).
FT   CONFLICT    414    414       R -> G (IN REF. 2 AND 3).
FT   CONFLICT    415    415       A -> L (IN REF. 2).
FT   CONFLICT    420    421       DH -> ER (IN REF. 3).
FT   CONFLICT    430    430       P -> S (IN REF. 1).
SQ   SEQUENCE   430 AA;  47715 MW;  FAEBB78FAF4723C3 CRC64;
```

```
  Query Match             60.9%;  Score 98;  DB 1;  Length 430;
  Best Local Similarity   64.5%;  Pred. No. 2e-06;
  Matches   20;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           | |:| :|||::|  |||||| ||||  ||| 
Db     281 LGHSIAQLNAQHNHIFYFKIDTVKKATSQVV 311


RESULT   8
KNT2_RAT
ID   KNT2_RAT       STANDARD;      PRT;   430 AA.
AC   P08932;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   T-kininogen II precursor (Major acute phase protein) (Alpha-1-MAP)
DE   (Thiostatin) [Contains: T-kinin].
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86008264; PubMed=2413018;
RA   Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT   "Primary structures of the mRNAs encoding the rat precursors for
RT   bradykinin and T-kinin. Structural relationship of kininogens with
RT   major acute phase protein and alpha 1-cysteine proteinase
RT   inhibitor.";
RL   J. Biol. Chem. 260:12054-12059(1985).
CC   -!- FUNCTION: KININOGENS ARE PLASMA GLYCOPROTEINS WITH A NUMBER OF
CC       FUNCTIONS: (1) AS PRECURSOR OF THE ACTIVE PEPTIDE BRADYKININ THEY
CC       EFFECT SMOOTH MUSCLE CONTRACTION, INDUCTION OF HYPOTENSION AND
CC       INCREASE OF VASCULAR PERMEABILITY. (2) THEY PLAY A ROLE IN BLOOD
CC       COAGULATION BY HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND
CC       FACTOR XI NEXT TO FACTOR XII. (3) THEY ARE INHIBITOR OF THIOL
CC       PROTEASES.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- INDUCTION: IN RESPONSE TO AN INFLAMMATORY STIMULANT. T-KININOGEN
CC       II SYNTHESIS IS INDUCED AND THE PLASMA CONCENTRATION OF
CC       T-KININOGEN I IS RAISED.
CC   -!- PTM: AS T-KININ IS PRECEDED BY A MET INSTEAD OF AN ARG OR LYS, IT
CC       IS NOT RELEASED FROM ITS PRECURSOR BY EITHER TISSUE OR PLASMA
CC       KALLIKREIN.
CC   -!- MISCELLANEOUS: RAT EXPRESS FOUR TYPES OF KININOGENS: THE CLASSICAL
CC       HMW AND LMW KININOGENS PRODUCED BY ALTERNATIVE SPLICING OF THE
CC       SAME GENE, AND TWO ADDITIONAL LMW-LIKE KININOGENS: T-I AND T-II.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M11885; AAA41491.1; -.
DR   PIR; B28055; B28055.
DR   GlycoSuiteDB; P08932; -.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   Pfam; PF00031; cystatin; 3.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Multigene family;
KW   Thiol protease inhibitor; Bradykinin; Acute phase; Signal.
FT   SIGNAL        1     18
FT   CHAIN        19    430       KININOGEN, T-II.
FT   CHAIN        19    375       HEAVY CHAIN.
FT   PEPTIDE     376    386       T-KININ.
FT   CHAIN       387    430       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    257       CYSTATIN-LIKE 2.
FT   DOMAIN      258    375       CYSTATIN-LIKE 3.
FT   DISULFID     28    404       INTERCHAIN (BY SIMILARITY).
FT   DISULFID     83     94       BY SIMILARITY.
FT   DISULFID    107    125       BY SIMILARITY.
FT   DISULFID    141    144       BY SIMILARITY.
FT   DISULFID    205    217       BY SIMILARITY.
FT   DISULFID    228    247       BY SIMILARITY.
FT   DISULFID    263    266       BY SIMILARITY.
FT   DISULFID    327    339       BY SIMILARITY.
FT   DISULFID    350    369       BY SIMILARITY.
FT   CARBOHYD     82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    326    326       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   430 AA;  47524 MW;  43EDF02D1BF55076 CRC64;


  Query Match             60.9%;  Score 98;  DB 1;  Length 430;
  Best Local Similarity   64.5%;  Pred. No. 2e-06;
  Matches   20;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           | |:| :|||::|  |||||| ||||  ||| 
Db     281 LGHSIAQLNAQHNHLFYFKIDTVKKATSQVV 311


RESULT   9
KNG_MOUSE
ID   KNG_MOUSE      STANDARD;      PRT;   661 AA.
AC   O08677; O08676;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Kininogen precursor [Contains: Bradykinin].
GN   KNG.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC   STRAIN=C57BL/6 x CBA; TISSUE=Liver;
RA   Takano M., Kondoh J., Yayama K., Okamoto H.;
RT   "Molecular cloning of cDNAs for mouse low- and high- molecular
RT   kininogen.";
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC       HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
CC       FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-
CC       INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
CC       BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
CC       PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
CC       CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
CC       DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
CC       MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
CC       PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
CC       OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
CC       A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
CC       INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
CC       KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
CC       CLOTTING (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: PLASMA.
```

```
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D84435; BAA19743.1; -.
DR   EMBL; D84415; BAA19742.1; -.
DR   MGD; MGI:1097705; Kng.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 1.
KW   Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW   Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW   Alternative splicing.
FT   SIGNAL       1     18       POTENTIAL.
FT   CHAIN       19    661       KININOGEN.
FT   CHAIN       19    379       KININOGEN HEAVY CHAIN.
FT   CHAIN      380    388       BRADYKININ.
FT   CHAIN      389    661       KININOGEN LIGHT CHAIN.
FT   DOMAIN      19    135       CYSTATIN-LIKE 1.
FT   DOMAIN     136    257       CYSTATIN-LIKE 2.
FT   DOMAIN     258    379       CYSTATIN-LIKE 3.
FT   DOMAIN     439    524       HIS-RICH.
FT   DISULFID    28    631       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    83     94       BY SIMILARITY.
FT   DISULFID   107    125       BY SIMILARITY.
FT   DISULFID   141    144       BY SIMILARITY.
FT   DISULFID   205    217       BY SIMILARITY.
FT   DISULFID   228    247       BY SIMILARITY.
FT   DISULFID   263    266       BY SIMILARITY.
FT   DISULFID   327    339       BY SIMILARITY.
FT   DISULFID   350    369       BY SIMILARITY.
FT   CARBOHYD    82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   168    168       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   204    204       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   242    242       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   401    432       VSPPYIAREQEERDAETEQGPTHGHGWLHEKQ -> RLLRA
FT                               CEYKGRLSKAGAEPAPERQAESSQVKQ (IN ISOFORM
FT                               LMW).
FT   VARSPLIC   433    661       MISSING (IN ISOFORM LMW).
SQ   SEQUENCE   661 AA;  73102 MW;  774460258D58796E CRC64;


  Query Match             60.9%;  Score 98;  DB 1;  Length 661;
  Best Local Similarity   64.5%;  Pred. No. 3.1e-06;
  Matches   20;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
           | |:| :|||||:  ||:||| ||||  |||
Db     281 LGHSIAQLNAENDHPFYYKIDTVKKATSQVV 311


RESULT  10
CYTF_MOUSE
ID   CYTF_MOUSE     STANDARD;      PRT;   144 AA.
AC   O89098;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Cystatin F precursor (Leukocystatin) (Cystatin 7) (Cystatin-like
DE   metastasis-associated protein) (CMAP).
GN   CST7.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98298157; PubMed=9632704;
RA   Halfon S., Ford J., Foster J., Dowling L., Lucian L., Sterling M.,
RA   Xu Y., Weiss M., Ikeda M., Liggett D., Helms A., Caux C., Lebecque S.,
RA   Hannum C., Menon S., McClanahan T., Gorman D., Zurawski G.;
RT   "Leukocystatin, a new class II cystatin expressed selectively by
RT   hematopoietic cells.";
RL   J. Biol. Chem. 273:16400-16408(1998).
CC   -!- FUNCTION: INHIBITS PAPAIN AND CATHEPSIN L BUT WITH AFFINITIES
CC       LOWER THAN OTHER CYSTATINS. MAY PLAY A ROLE IN IMMUNE REGULATION
CC       THROUGH INHIBITION OF A UNIQUE TARGET IN THE HEMATOPOIETIC SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC   -!- SIMILARITY: BELONGS TO THE CYSTATIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF031826; AAC40140.1; -.
DR   EMBL; AF031825; AAC40139.1; -.
DR   HSSP; P01038; 1A90.
DR   MGD; MGI:1298217; Cst7.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   Pfam; PF00031; cystatin; 1.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00287; CYSTATIN; FALSE_NEG.
KW   Thiol protease inhibitor; Glycoprotein; Signal.
FT   SIGNAL       1     18       POTENTIAL.
FT   CHAIN       19    144       CYSTATIN F.
FT   ACT_SITE    36     36       REACTIVE SITE.
FT   SITE        80     84       SECONDARY AREA OF CONTACT.
FT   DISULFID    98    109       BY SIMILARITY.
FT   DISULFID   123    143       BY SIMILARITY.
SQ   SEQUENCE   144 AA;  16380 MW;  B5837334C1B4A89C CRC64;


  Query Match             37.9%;  Score 61;  DB 1;  Length 144;
  Best Local Similarity   48.3%;  Pred. No. 0.094;
  Matches   14;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       4 HTITKLNAENNATFYFKIDNVKKARVQVV 32
           |:: | |   |  | ||  :| || ||||
Db      54 HSVEKFNNCTNDIFLFKESHVSKALVQVV 82


RESULT  11
UFD1_DROME
ID   UFD1_DROME     STANDARD;      PRT;   316 AA.
AC   Q9VTF9;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ubiquitin fusion degradation protein 1 homolog (UB fusion protein 1).
GN   CG6233.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
```

```
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: FUNCTIONS AT A POST-UBIQUITATION STEP IN THE UBIQUITIN
CC       FUSION DEGRADATION (UFD) PATHWAY (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE UFD1 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE003546; AAF50090.1; -.
DR   FlyBase; FBgn0036136; CG6233.
KW   Ubiquitin conjugation.
SQ   SEQUENCE   316 AA;  34791 MW;  B7C1EA88CCA5CDA3 CRC64;


  Query Match             32.3%;  Score 52;  DB 1;  Length 316;
  Best Local Similarity   50.0%;  Pred. No. 4;
  Matches   12;  Conservative    4;  Mismatches    6;  Indels    2;  Gaps    1;

Qy       5 TITKLNAENNATFYFKIDNVKKAR 28
           |:|:|| |      ||: ||||:|
Db      52 TLTRLNVE--YPMLFKLTNVKKSR 73


RESULT  12
HS83_LEIAM
ID   HS83_LEIAM     STANDARD;      PRT;   701 AA.
AC   P27741;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-OCT-1993 (Rel. 27, Last annotation update)
DE   Heat shock protein 83 (HSP 83).
GN   HSP83.
OS   Leishmania amazonensis.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5659;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ISOLATE MHOM/BR/77/LTB0016;
RX   MEDLINE=91101661; PubMed=2270107;
RA   Shapria M., Pedraza G.;
RT   "Sequence analysis and transcriptional activation of heat shock
RT   protein 83 of Leishmania mexicana amazonensis.";
RL   Mol. Biochem. Parasitol. 42:247-256(1990).
CC   -!- FUNCTION: MOLECULAR CHAPERONE. HAS ATPASE ACTIVITY
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M92926; AAA29250.1; -.
DR   PIR; A44943; A44943.
DR   HSSP; P07900; 1YER.
DR   InterPro; IPR003594; HATPase_c.
DR   InterPro; IPR001404; HSP90.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00183; HSP90; 1.
DR   PRINTS; PR00775; HEATSHOCK90.
DR   SMART; SM00387; HATPase_c; 1.
DR   PROSITE; PS00298; HSP90; 1.
KW   Chaperone; ATP-binding; Heat shock.
SQ   SEQUENCE   701 AA;  80584 MW;  05100A1324A80812 CRC64;


  Query Match             32.3%;  Score 52;  DB 1;  Length 701;
  Best Local Similarity   38.5%;  Pred. No. 9.4;
  Matches   10;  Conservative    6;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAENNATFYFKIDNVKK 26
             ||   :|:: || |: :|   |: ||
Db     448 TLKDYVTRMKAEQNSIYYITGDSKKK 473


RESULT  13
YDS4_SCHPO
ID   YDS4_SCHPO     STANDARD;      PRT;   306 AA.
AC   O14180;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Hypothetical 35.8 kDa protein C4F8.04 in chromosome I.
GN   SPAC4F8.04.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE IMP4 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Z98530; CAB11051.1; -.
DR   InterPro; IPR002799; IMP4.
DR   Pfam; PF01945; IMP4; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   306 AA;  35835 MW;  2029FBB2358EC01E CRC64;


  Query Match             31.1%;  Score 50;  DB 1;  Length 306;
  Best Local Similarity   30.8%;  Pred. No. 7.4;
  Matches    8;  Conservative    9;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFYFKIDNVKKAR 28
           |: :| ::  |  :||| : |:: |:
Db     158 TNALTLVHLPNGPSFYFTLSNLQTAK 183


RESULT  14
A2HS_RAT
ID   A2HS_RAT       STANDARD;      PRT;   352 AA.
AC   P24090;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Alpha-2-HS-glycoprotein precursor (Fetuin-A) (Glycoprotein PP63)
DE   (59 kDa bone sialic acid-containing protein) (BSP).
GN   AHSG OR FETUA.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92174906; PubMed=1371750;
RA   Rauth G., Poeschke O., Fink E., Eulitz M., Tippmer S., Kellerer M.,
RA   Haering H., Nawratil P., Haasemann M., Jahnen-Dechent W.,
RA   Mueller-Esterl W.;
RT   "The nucleotide and partial amino acid sequences of rat fetuin.
RT   Identity with the natural tyrosine kinase inhibitor of the rat
RT   insulin receptor.";
RL   Eur. J. Biochem. 204:523-529(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=89354538; PubMed=2766355;
RA   Auberger P., Falquerho L., Contreres J.O., Pages G., le Cam G.,
RA   Rossi B., le Cam A.;
RT   "Characterization of a natural inhibitor of the insulin receptor
RT   tyrosine kinase: cDNA cloning, purification, and anti-mitogenic
RT   activity.";
RL   Cell 58:631-640(1989).
RN   [3]
RP   REVISIONS, SEQUENCE FROM N.A.
RX   MEDLINE=91200667; PubMed=1849862;
RA   Falquerho L., Patey G., Paqureau L., Rossi V., Lahuna O., Szpirer J.,
RA   Szpirer C., Levan G., le Cam A.;
RT   "Primary structure of the rat gene encoding an inhibitor of the
RT   insulin receptor tyrosine kinase.";
RL   Gene 98:209-216(1991).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93206661; PubMed=7681247;
RA   Ohnishi T., Nakamura O., Ozawa M., Arakaki N., Muramatsu T.,
RA   Daikuhara Y.;
RT   "Molecular cloning and sequence analysis of cDNA for a 59 kD bone
RT   sialoprotein of the rat: demonstration that it is a counterpart of
RT   human alpha 2-HS glycoprotein and bovine fetuin.";
RL   J. Bone Miner. Res. 8:367-377(1993).
RN   [5]
RP   IDENTITY OF PP63 WITH FETUIN.
RX   MEDLINE=91190111; PubMed=1707273;
RA   Haasemann M., Nawratil P., Mueller-Esterl W.;
RT   "Rat tyrosine kinase inhibitor shows sequence similarity to human
RT   alpha 2-HS glycoprotein and bovine fetuin.";
RL   Biochem. J. 274:899-902(1991).
RN   [6]
RP   IDENTITY OF PP63 WITH FETUIN.
RX   MEDLINE=92119718; PubMed=1370655;
RA   Brown W.M., Christie D.L., Dziegielewska K.M., Saunders N.R., Yang F.;
RT   "The rat protein encoded by clone pp63 is a fetuin/alpha 2-HS
RT   glycoprotein-like molecule, but is it the tyrosine kinase inhibitor
RT   pp63?";
RL   Cell 68:7-8(1992).
CC   -!- FUNCTION: COULD INHIBIT BOTH INSULIN-RECEPTOR TYROSINE KINASE
CC       ACTIVITY AND INSULIN-STIMULATED RECEPTOR AUTOPHOSPHORYLATION AND,
CC       CONCOMITANTLY, ANTAGONIZE THE MITOGENIC EFFECT OF THE HORMONE IN
CC       CULTURED RAT HEPATOMA CELLS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: SYNTHESIZED IN LIVER AND SECRETED BY THE
CC       HEPATOCYTES IN THE BLOOD.
CC   -!- PTM: UNDERGOES COMPLEX POSTTRANSLATIONAL MODIFICATION INVOLVING
CC       N-GLYCOSYLATION, AND ADDITION OF FUCOSE AND SIALIC ACID RESIDUES.
CC       PHOSPHORYLATION OCCURS AT A SERINE RESIDUE.
CC   -!- SIMILARITY: BELONGS TO THE FETUIN FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 CYSTATIN-LIKE DOMAINS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M29758; AAA75502.1; -.
DR   EMBL; X63446; CAA45042.1; -.
DR   EMBL; D10261; BAA01101.1; -.
DR   PIR; A32827; A32827.
DR   PIR; S22382; S22382.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR001363; Fetuin.
DR   Pfam; PF00031; cystatin; 2.
DR   SMART; SM00043; CY; 2.
DR   PROSITE; PS01254; FETUIN_1; 1.
DR   PROSITE; PS01255; FETUIN_2; 1.
KW   Glycoprotein; Signal; Repeat; Phosphorylation.
FT   SIGNAL       1     18       POTENTIAL.
FT   CHAIN       19    352       ALPHA-2-HS-GLYCOPROTEIN.
FT   DOMAIN      27    144       CYSTATIN-LIKE 1.
FT   DOMAIN     145    255       CYSTATIN-LIKE 2.
FT   DISULFID    32    343       BY SIMILARITY.
FT   DISULFID    89    100       BY SIMILARITY.
FT   DISULFID   114    132       BY SIMILARITY.
FT   DISULFID   146    149       BY SIMILARITY.
FT   DISULFID   208    219       BY SIMILARITY.
FT   DISULFID   230    247       BY SIMILARITY.
FT   CARBOHYD    99     99       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   156    156       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   176    176       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   SITE       143    144       CLEAVAGE (BY TRYPSIN) (POTENTIAL).
FT   CONFLICT    46     46       D -> H (IN REF. 2 AND 3).
FT   CONFLICT    76     76       E -> Q (IN REF. 2 AND 3).
SQ   SEQUENCE   352 AA;  37982 MW;  43564F60F3C7C90A CRC64;


  Query Match             30.7%;  Score 49.5;  DB 1;  Length 352;
  Best Local Similarity   35.7%;  Pred. No. 10;
  Matches   10;  Conservative    7;  Mismatches   10;  Indels    1;  Gaps    1;

Qy       1 TLTHTITKLNAENNATFYFKIDNVKKAR 28
           |:   :   ||:|| | |||:  : :|:
Db     163 TVKTALAAFNAQNNGT-YFKLVEISRAQ 189
```

```
RESULT 15
T2H2_HAEPA
ID   T2H2_HAEPA     STANDARD;      PRT;   358 AA.
AC   P36433;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Type II restriction enzyme HpaII (EC 3.1.21.4) (Endonuclease HpaII)
DE   (R.HpaII).
GN   HPAIIR.
OS   Haemophilus parainfluenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=729;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 4969;
RX   MEDLINE=94237495; PubMed=7514149;
RA   Kulakauskas S., Barsomian J.M., Lubys A., Roberts R.J., Wilson G.G.;
RT   "Organization and sequence of the HpaII restriction-modification
RT   system and adjacent genes.";
RL   Gene 142:9-15(1994).
CC   -!- FUNCTION: RECOGNIZES THE DOUBLE-STRANDED SEQUENCE
CC       CCGG AND CLEAVES AFTER C-1.
CC   -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage of DNA to give
CC       specific double-stranded fragments with terminal 5'-phosphates.
CC   -!- SUBUNIT: HOMODIMER.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L17342; AAA20482.1; -.
DR   REBASE; 1159; HpaII.
KW   Hydrolase; Endonuclease; Nuclease; Restriction system.
SQ   SEQUENCE   358 AA;  40925 MW;  ECB312180C80303E CRC64;


  Query Match             30.4%;  Score 49;  DB 1;  Length 358;
  Best Local Similarity   38.5%;  Pred. No. 12;
  Matches   10;  Conservative    5;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       7 TKLNAENNATFYFKIDNVKKARVQVV 32
           | |||     | ||| |:|  ::: :
Db     157 TLLNASKATNFTFKIYNLKDKQIEYI 182


Search completed: July  1, 2002, 16:30:13
Job time: 622 sec
```

GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:20:36 ; Search time 46.58 Seconds
(without alignments)
24.755 Million cell updates/sec

Title: US-09-461-061A-2
Perfect score: 60
Sequence: 1 TLTHTITKLNAE 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 60 | 100.0 | 427 | 1 | KGHUL1 | kininogen, LMW pre |
| 2 | 60 | 100.0 | 644 | 1 | KGHUH1 | kininogen, HMW pre |
| 3 | 42 | 70.0 | 436 | 1 | KGBOL1 | kininogen, LMW I p |
| 4 | 42 | 70.0 | 621 | 1 | KGBOH1 | kininogen, HMW I p |
| 5 | 41 | 68.3 | 434 | 1 | KGBOL2 | kininogen, LMW II |
| 6 | 41 | 68.3 | 617 | 2 | I56530 | gene VGF protein - |
| 7 | 41 | 68.3 | 619 | 1 | KGBOH2 | kininogen, HMW II |
| 8 | 41 | 68.3 | 711 | 2 | S05381 | VGF8a protein prec |
| 9 | 39 | 65.0 | 534 | 2 | F82615 | methyltransferase |
| 10 | 38 | 63.3 | 498 | 2 | S49776 | hypothetical prote |
| 11 | 37 | 61.7 | 90 | 2 | A60526 | complement C3 - ax |
| 12 | 37 | 61.7 | 103 | 2 | T03013 | hypothetical prote |
| 13 | 37 | 61.7 | 174 | 2 | AI0618 | conserved hypothet |
| 14 | 37 | 61.7 | 786 | 2 | T23883 | hypothetical prote |
| 15 | 37 | 61.7 | 1992 | 1 | S02771 | myosin heavy chain |
| 16 | 36 | 60.0 | 197 | 2 | C84084 | hypothetical prote |
| 17 | 36 | 60.0 | 305 | 2 | H82888 | methionyl-tRNA for |
| 18 | 36 | 60.0 | 371 | 2 | AE1348 | a probable compete |
| 19 | 36 | 60.0 | 371 | 2 | AI1718 | probable competenc |
| 20 | 36 | 60.0 | 395 | 2 | S75952 | hypothetical prote |
| 21 | 36 | 60.0 | 1231 | 2 | A54803 | microtubule-associ |
| 22 | 35.5 | 59.2 | 269 | 2 | S63631 | acid proteinase ea |
| 23 | 35 | 58.3 | 157 | 2 | D82209 | probable antioxida |
| 24 | 35 | 58.3 | 178 | 2 | H82674 | heat shock protein |
| 25 | 35 | 58.3 | 184 | 2 | A61196 | genome polyprotein |
| 26 | 35 | 58.3 | 194 | 2 | S06067 | nonstructural prot |
| 27 | 35 | 58.3 | 202 | 2 | T46515 | probable two-compo |
| 28 | 35 | 58.3 | 288 | 2 | A83443 | probable transcrip |
| 29 | 35 | 58.3 | 433 | 2 | A28055 | K-kininogen, LMW I |
| 30 | 35 | 58.3 | 555 | 2 | D86732 | DNA repair protein |
| 31 | 35 | 58.3 | 589 | 2 | C82279 | hypothetical prote |
| 32 | 35 | 58.3 | 613 | 2 | T27077 | hypothetical prote |
| 33 | 35 | 58.3 | 639 | 2 | A25486 | kininogen, HMW I p |
| 34 | 35 | 58.3 | 716 | 2 | JQ1366 | polyprotein - hepa |
| 35 | 35 | 58.3 | 781 | 2 | F86457 | unknown protein, 3 |
| 36 | 35 | 58.3 | 876 | 2 | PC2219 | polypeptide - hepa |
| 37 | 35 | 58.3 | 890 | 2 | A97750 | DNA mismatch repai |
| 38 | 35 | 58.3 | 1211 | 2 | S54500 | alpha,alpha-trehal |
| 39 | 35 | 58.3 | 2551 | 2 | B98047 | hypothetical prote |
| 40 | 35 | 58.3 | 3010 | 1 | GNWVTC | genome polyprotein |
| 41 | 35 | 58.3 | 3010 | 1 | GNWVCJ | genome polyprotein |
| 42 | 35 | 58.3 | 3010 | 1 | A45573 | genome polyprotein |
| 43 | 35 | 58.3 | 3010 | 1 | S18030 | genome polyprotein |
| 44 | 35 | 58.3 | 3010 | 1 | GNWVTW | genome polyprotein |
| 45 | 35 | 58.3 | 3011 | 1 | GNWVCH | genome polyprotein |

ALIGNMENTS

RESULT 1
KGHUL1
kininogen, LMW precursor [validated] - human
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Homo sapiens (man)
C;Date: 06-Jul-1982 #sequence_revision 27-Nov-1985 #text_change 08-Dec-2000
C;Accession: A01280; B25276; A27900; A27699; A31905; A34030
R;Ohkubo, I.; Kurachi, K.; Takasawa, T.; Shiokawa, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A;Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its ide
A;Reference number: A90490; MUID:85122621
A;Accession: A01280
A;Molecule type: mRNA
A;Residues: 1-427 <OHK>
A;Cross-references: GB:K02566; NID:g177889; PIDN:AAA35497.1; PID:g177890
R;Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A;Title: Cloning and sequence analysis of cDNAs for human high molecular weight and l
A;Reference number: A92544; MUID:85234582
A;Accession: B25276
A;Molecule type: mRNA
A;Residues: 1-427 <TAK>
A;Cross-references: GB:M11437; NID:g186751; PIDN:AAB59551.1; PID:g386853
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Rauth, G.; Mueller-Esterl, W.
in Kinins IV, part A, Greenbaum, L.M., and Margolius, H.S., eds., pp.91-95, Plenum, N
A;Title: Amino acid sequence of the light chain of human low molecular mass kininogen
A;Reference number: A27900
A;Accession: A27900
A;Molecule type: protein
A;Residues: 390-427 <LOT>
R;Mindroiu, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
Biochem. Biophys. Res. Commun. 152, 519-526, 1988
A;Title: A new kinin moiety in human plasma kininogens.
A;Reference number: A27699; MUID:88209021
A;Accession: A27699
A;Molecule type: protein
A;Residues: 380-389 <MIN>
R;Maeda, H.; Matsumura, Y.; Kato, H.
J. Biol. Chem. 263, 16051-16054, 1988
A;Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic f
A;Reference number: A31905; MUID:89034061
A;Accession: A31905
A;Molecule type: protein
A;Residues: 381-389 <MAE>
R;Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
Biochem. Biophys. Res. Commun. 150, 511-516, 1988
A;Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human p
A;Reference number: A34030; MUID:88106632
A;Accession: A34030
A;Molecule type: protein
A;Residues: 380-389 <SAS>

R;Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 260, 8610-8617, 1985
A;Title: Structural organization of the human kininogen gene and a model for its evoluti
A;Reference number: A92545; MUID:85234583
A;Contents: annotation; gene organization
R;Pierce, J.V.
Fed. Proc. 27, 52-57, 1968
A;Title: Structural features of plasma kinins and kininogens.
A;Reference number: A91455; MUID:90255622
A;Contents: annotation; bradykinin
C;Comment: The LMW kininogen precursor is produced from the same gene as the HMW form (s
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Genetics:
A;Gene: GDB:KNG
A;Cross-references: GDB:125256; OMIM:228960
A;Map position: 3q27-3q27
A;Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3; 401/3
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; glyc
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-427/Product: LMW prokininogen (kininogen I) #status predicted <MAT>
F;19-389,390-427/Product: LMW kininogen II #status predicted <MAT2>
F;19-379/Product: LMW kininogen heavy chain #status predicted <HCH>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>
F;380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
F;390-427/Product: LMW kininogen light chain #status experimental <LCH>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;28-407,83-94,107-126,142-145,206-218,229-248,264-267,328-340,351-370/Disulfide bonds:
F;48,169,205,294/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
F;383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F;389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;401/Binding site: carbohydrate (Thr) (covalent) #status absent

Query Match 100.0%; Score 60; DB 1; Length 427;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNAE 12
||||||||||||
Db 281 TLTHTITKLNAE 292

RESULT 2
KGHUH1
kininogen, HMW precursor [validated] - human
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen; prokininogen
N;Contains: bradykinin (kallidin I); HMW kininogen I; HMW kininogen II; low molecular we
C;Species: Homo sapiens (man)
C;Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 08-Dec-2000
C;Accession: A01279; A25276; S32422; A91153; A24871; A27899; A27699; A31905; A34030; S02
R;Ohkubo, I.; Kurachi, K.; Takasawa, T.; Shiokawa, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A;Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its identi
A;Reference number: A90490; MUID:85122621
A;Accession: A01279
A;Molecule type: mRNA
A;Residues: 1-389 <OHK>
A;Cross-references: GB:K02566; NID:g177889
R;Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A;Title: Cloning and sequence analysis of cDNAs for human high molecular weight and low
A;Reference number: A92544; MUID:85234582
A;Accession: A25276
A;Molecule type: mRNA
A;Residues: 1-592,'I',594-644 <TAK>
A;Cross-references: GB:M11437; NID:g186751; PIDN:AAB59550.1; PID:g386852
R;Auerswald, E.A.; Roessler, D.; Mentele, R.; Assfalg-Machleidt, I.
FEBS Lett. 321, 93-97, 1993
A;Title: Cloning, expression and characterization of human kininogen domain 3.
A;Reference number: S32422; MUID:93223854
A;Accession: S32422
A;Molecule type: mRNA
A;Residues: 'ANSM',253-377 <AUE>
A;Note: differences are due to known cloning artifacts
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A;Title: The amino acid sequence of the light chain of human high-molecular-mass kini
A;Reference number: A91153; MUID:86030270
A;Accession: A91153
A;Molecule type: protein
A;Residues: 379-644 <LOT>
A;Note: the bradykinin sequence preceding the light chain sequence was not determined
R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
Eur. J. Biochem. 154, 471-478, 1986
A;Title: Completion of the primary structure of human high-molecular-mass kininogen.
A;Reference number: A24871; MUID:86108361
A;Accession: A24871
A;Molecule type: protein
A;Residues: 'Z',20-380 <KEL1>
R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
in Kinins IV, Greenbaum, L.M., and Margolius, H.S., ed., pp.85-89, Plenum Press, New
A;Title: Amino acid sequence of the light chain of human high molecular mass kininoge
A;Reference number: A27899
A;Accession: A27899
A;Molecule type: protein
A;Residues: 379-389,'K',390-407,'Q',409-644 <KEL2>
R;Mindroiu, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
Biochem. Biophys. Res. Commun. 152, 519-526, 1988
A;Title: A new kinin moiety in human plasma kininogens.
A;Reference number: A27699; MUID:88209021
A;Accession: A27699
A;Molecule type: protein
A;Residues: 380-389 <MIN>
R;Maeda, H.; Matsumura, Y.; Kato, H.
J. Biol. Chem. 263, 16051-16054, 1988
A;Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic f
A;Reference number: A31905; MUID:89034061
A;Accession: A31905
A;Molecule type: protein
A;Residues: 381-389 <MAE>
R;Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
Biochem. Biophys. Res. Commun. 150, 511-516, 1988
A;Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human p
A;Reference number: A34030; MUID:88106632
A;Accession: A34030
A;Molecule type: protein
A;Residues: 380-389 <SAS>
R;Lenarcic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
Biol. Chem. Hoppe-Seyler 369, 257-261, 1988
A;Title: Human cathepsin B and cysteine proteinase inhibitors (CPIs) in inflammatory
A;Reference number: S02482; MUID:89076517
A;Accession: S02482
A;Molecule type: protein
A;Residues: 1-19;189-192;310-314;381-389 <LEN1>
R;Kato, H.; Matsumura, Y.; Maeda, H.
FEBS Lett. 232, 252-254, 1988
A;Title: Isolation and identification of hydroxyproline analogues of bradykinin in hu
A;Reference number: A61495; MUID:88211869
A;Accession: A61495
A;Molecule type: protein
A;Residues: 380-389 <KAT1>
A;Experimental source: urine
A;Note: this peptide had Pro-383 modified to 4-hydroxyproline
A;Accession: B61495
A;Molecule type: protein
A;Residues: 381-389 <KAT2>
A;Experimental source: urine
A;Note: this peptide had Pro-383 modified to 4-hydroxyproline
A;Accession: C61495

A;Molecule type: protein
A;Residues: 380-389 <KAT3>
R;Lenarcic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
FEBS Lett. 280, 211-215, 1991
A;Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
A;Reference number: S14303; MUID:91192133
A;Accession: S14447
A;Molecule type: protein
A;Residues: 264-359,'N',361-375 <LEN2>
R;Little, S.S.; Johnson, D.A.
Biochem. J. 307, 341-346, 1995
A;Title: Human mast cell tryptase isoforms: separation and examination of substrate-spec
A;Reference number: S55239; MUID:95251593
A;Accession: S55239
A;Molecule type: protein
A;Residues: 450-452,'X',454,'X',456 <LIT>
R;Straczek, J.; Maachi, F.; le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabet, P.; Bellevil
FEBS Lett. 373, 207-211, 1995
A;Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like
A;Reference number: S68059; MUID:96033974
A;Accession: S68059
A;Molecule type: protein
A;Residues: 431-434 <STR>
R;Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 260, 8610-8617, 1985
A;Title: Structural organization of the human kininogen gene and a model for its evoluti
A;Reference number: A92545; MUID:85234583
A;Contents: annotation; gene organization
R;Pierce, J.V.
Fed. Proc. 27, 52-57, 1968
A;Title: Structural features of plasma kinins and kininogens.
A;Reference number: A91455; MUID:90255622
A;Contents: annotation; bradykinin
C;Comment: The HMW kininogen precursor and the LMW form are produced from the same gene
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Genetics:
A;Gene: GDB:KNG
A;Cross-references: GDB:125256; OMIM:228960
A;Map position: 3q27-3q27
A;Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>
F;19-379,390-644/Product: HMW kininogen II #status experimental <MAT2>
F;19-379/Domain: HMW kininogen heavy chain #status experimental <HCH>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>
F;380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
F;390-644/Domain: HMW kininogen light chain #status experimental <LCH>
F;421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
F;431-434/Product: low molecular weight growth promoting factor #status experimental <GF
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;28-614,83-94,107-126,142-145,206-218,229-248,264-267,328-340,351-370/Disulfide bonds:
F;48/Binding site: carbohydrate (Asn) (covalent) #status absent
F;169,205,294/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
F;383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F;389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;401,533,542,546,557,571,593,628/Binding site: carbohydrate (Thr) (covalent) #status ex
F;577/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 60; DB 1; Length 644;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNAE 12
||||||||||||
Db 281 TLTHTITKLNAE 292

RESULT 3
KGBOL1
kininogen, LMW I precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
C;Accession: A01283
R;Nawa, H.; Kitamura, N.; Hirose, T.; Asai, M.; Inayama, S.; Nakanishi, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 90-94, 1983
A;Title: Primary structures of bovine liver low molecular weight kininogen precursors
A;Reference number: A93984; MUID:83117859
A;Accession: A01283
A;Molecule type: mRNA
A;Residues: 1-436 <NAW>
A;Cross-references: GB:J00010; GB:V00426; NID:g163256; PIDN:AAA30604.1; PID:g163257
C;Comment: The LMW kininogen precursor is produced from the same gene as the HMW form
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; g
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-436/Product: LMW kininogen I #status predicted <MAT>
F;19-378/Product: LMW kininogen I heavy chain #status predicted <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;379-388/Product: lysyl-bradykinin (kallidin II) #status predicted <KBDY>
F;380-388/Product: bradykinin (kallidin I) #status predicted <BDY>
F;389-436/Product: LMW kininogen I light chain #status experimental <LCH>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F;27-406,82-93,106-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bond
F;47,87,168,169,197,204/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;378-379/Cleavage site: Met-Lys (kallikrein) #status predicted
F;382/Modified site: 4-hydroxyproline (Pro) #status predicted
F;388-389/Cleavage site: Arg-Ser (kallikrein) #status predicted

Query Match 70.0%; Score 42; DB 1; Length 436;
Best Local Similarity 72.7%; Pred. No. 4;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LTHTITKLNAE 12
|:|:| |||||
Db 281 LSHSIAKLNAE 291

RESULT 4
KGBOH1
kininogen, HMW I precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
C;Accession: A01281; A91923; A91938; A29559
R;Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
Nature 305, 545-549, 1983
A;Title: A single gene for bovine high molecular weight and low molecular weight kini
A;Reference number: A93317; MUID:84014106
A;Accession: A01281
A;Molecule type: mRNA
A;Residues: 1-621 <KIT>
A;Cross-references: GB:V01491; GB:K01757; NID:g491; PIDN:CAA24735.1; PID:g492
R;Kato, H.; Nagasawa, S.; Suzuki, T.
J. Biochem. 67, 313-323, 1970
A;Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds a
A;Reference number: A91923; MUID:70180420
A;Accession: A91923

A;Molecule type: protein
A;Residues: 378-393 <KAT>
R;Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
J. Biochem. 77, 55-68, 1975
A;Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Ami
A;Reference number: A91938; MUID:75170265
A;Accession: A91938
A;Molecule type: protein
A;Residues: 458-498 <HAN>
R;Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga,
J. Biol. Chem. 262, 2768-2779, 1987
A;Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of c
A;Reference number: A92627; MUID:87137530
A;Accession: A29559
A;Molecule type: protein
A;Residues: 'Z',20-123,'I',125-127,'I',129-378 <SUE>
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A;Title: The amino acid sequence of the light chain of human high-molecular-mass kininog
A;Reference number: A91153; MUID:86030270
A;Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
R;Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
Seikagaku 56, 808, 1984
A;Title: Disulfide bonds in bovine HMW kininogens.
A;Reference number: A94300
A;Contents: annotation; disulfide bonds
A;Note: article in Japanese
C;Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-621/Product: HMW prokininogen I #status predicted <MAT>
F;19-379/Product: HMW kininogen I heavy chain #status experimental <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;379-388/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;380-388/Product: bradykinin (kallidin I) #status experimental <BDY>
F;389-621/Product: HMW kininogen I light chain #status experimental <LCH>
F;417-488/Region: glycine/histidine/lysine-rich
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;27-591,82-93,106-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bonds:
F;87,168,169,204/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
F;197/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F;378-379/Cleavage site: Met-Lys (kallikrein) #status experimental
F;382/Modified site: 4-hydroxyproline (Pro) #status predicted
F;388-389/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;398,406,512/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;399,400,520,524,536,548,553,570/Binding site: carbohydrate (Thr) (covalent) #status ex
F;498-499/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match 70.0%; Score 42; DB 1; Length 621;
Best Local Similarity 72.7%; Pred. No. 5.9;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LTHTITKLNAE 12
|:|:| |||||
Db 281 LSHSIAKLNAE 291

RESULT 5
KGBOL2
kininogen, LMW II precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 28-May-1999
C;Accession: A01284
R;Nawa, H.; Kitamura, N.; Hirose, T.; Asai, M.; Inayama, S.; Nakanishi, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 90-94, 1983
A;Title: Primary structures of bovine liver low molecular weight kininogen precursors
A;Reference number: A93984; MUID:83117859
A;Accession: A01284
A;Molecule type: mRNA
A;Residues: 1-434 <NAW>
A;Cross-references: GB:V00427; GB:J00011; NID:g489; PIDN:CAA23710.1; PID:g490
C;Comment: The LMW kininogen precursor is produced from the same gene as the HMW form
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; g
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-434/Product: LMW kininogen II #status predicted <MAT>
F;19-377/Product: LMW kininogen I heavy chain #status predicted <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;261-372/Domain: cystatin homology <CY3>
F;377-386/Product: lysyl-bradykinin (kallidin II) #status predicted <KBDY>
F;378-386/Product: bradykinin (kallidin I) #status predicted <BDY>
F;387-434/Product: LMW kininogen I light chain #status experimental <LCH>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F;27-404,82-93,106-125,141-144,205-217,228-247,261-264,325-337,348-367/Disulfide bond
F;47,87,168,169,197,204,280/Binding site: carbohydrate (Asn) (covalent) #status predi
F;376-377/Cleavage site: Met-Lys (kallikrein) #status predicted
F;380/Modified site: 4-hydroxyproline (Pro) #status predicted
F;386-387/Cleavage site: Arg-Ser (kallikrein) #status predicted

Query Match 68.3%; Score 41; DB 1; Length 434;
Best Local Similarity 72.7%; Pred. No. 6.3;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LTHTITKLNAE 12
|:|:| |||||
Db 279 LNHSIAKLNAE 289

RESULT 6
I56530
gene VGF protein - rat
N;Alternate names: VGF
C;Species: Rattus norvegicus (Norway rat)
C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 01-Dec-2000
C;Accession: I56530; A39748; JH0454
R;Hawley, R.J.; Scheibe, R.J.; Wagner, J.A.
J. Neurosci. 12, 2573-2581, 1992
A;Title: NGF induces the expression of the VGF gene through a cAMP or response elemen
A;Reference number: I56530; MUID:92309005
A;Accession: I56530
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-617 <RES>
A;Cross-references: GB:M74223; NID:g207650; PIDN:AAA42336.1; PID:g207651
R;Salton, S.R.J.; Fischberg, D.J.; Dong, K.W.
Mol. Cell. Biol. 11, 2335-2349, 1991
A;Title: Structure of the gene encoding VGF, a nervous system-specific mRNA that is r
A;Reference number: A39748; MUID:91203852
A;Accession: A39748
A;Molecule type: DNA
A;Residues: 1-175,'D',177-617 <SAL1>
A;Cross-references: GB:M60522
R;Salton, S.R.J.
J. Neurochem. 57, 991-996, 1991
A;Title: Nucleotide sequence and regulatory studies of VGF, a nervous system-specific
A;Reference number: JH0454; MUID:91318308
A;Accession: JH0454
A;Molecule type: mRNA
A;Residues: 1-175,'D',177-617 <SAL2>
A;Cross-references: GB:M60525

A;Experimental source: cultured cell PC12, clone NGF33.1
C;Comment: This protein is induced in PC12 cells to a greater extent by nerve growth fac
C;Genetics:
A;Gene: VGF
C;Keywords: growth factor

Query Match 68.3%; Score 41; DB 2; Length 617;
Best Local Similarity 58.3%; Pred. No. 9.1;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 TLTHTITKLNAE 12
          | |||:|::| |
Db    190 TRTHTLTRVNLE 201
```

RESULT 7
KGBOH2
kininogen, HMW II precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
C;Accession: A01282; A91923; A91941; A91938; B29559
R;Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
Nature 305, 545-549, 1983
A;Title: A single gene for bovine high molecular weight and low molecular weight kininog
A;Reference number: A93317; MUID:84014106
A;Accession: A01282
A;Molecule type: mRNA
A;Residues: 1-619 <KIT>
A;Cross-references: GB:V01492; GB:K01758; NID:g493; PIDN:CAA24736.1; PID:g494
R;Kato, H.; Nagasawa, S.; Suzuki, T.
J. Biochem. 67, 313-323, 1970
A;Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds and
A;Reference number: A91923; MUID:70180420
A;Accession: A91923
A;Molecule type: protein
A;Residues: 376-391 <KAT>
R;Han, Y.N.; Kato, H.; Iwanaga, S.; Suzuki, T.
J. Biochem. 79, 1201-1222, 1976
A;Title: Primary structure of bovine plasma high-molecular-weight kininogen. The amino a
A;Reference number: A91941; MUID:76260155
A;Accession: A91941
A;Molecule type: protein
A;Residues: 387-455 <HAN>
A;Note: 398-Pro, 401-Val, and 455-Lys were also found
R;Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
J. Biochem. 77, 55-68, 1975
A;Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Ami
A;Reference number: A91938; MUID:75170265
A;Accession: A91938
A;Molecule type: protein
A;Residues: 456-496 <HA2>
R;Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga,
J. Biol. Chem. 262, 2768-2779, 1987
A;Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of c
A;Reference number: A92627; MUID:87137530
A;Accession: B29559
A;Molecule type: protein
A;Residues: 'Z',20-104,'E',106-256,'XX',257-376 <SUE>
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A;Title: The amino acid sequence of the light chain of human high-molecular-mass kininog
A;Reference number: A91153; MUID:86030270
A;Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
R;Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
Seikagaku 56, 808, 1984
A;Title: Disulfide bonds in bovine HMW kininogens.
A;Reference number: A94300
A;Contents: annotation; disulfide bonds
A;Note: article in Japanese
C;Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of
C;Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is i
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-619/Product: HMW kininogen II #status predicted <MAT>
F;19-376/Product: HMW kininogen II heavy chain #status experimental <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;261-372/Domain: cystatin homology <CY3>
F;377-386/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;378-386/Product: bradykinin (kallidin I) #status experimental <BDY>
F;387-619/Product: HMW kininogen II light chain #status experimental <LCH>
F;418-488/Region: glycine/histidine/lysine-rich
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
F;27-589,82-93,106-125,141-144,205-217,228-247,261-264,325-337,348-367/Disulfide bond
F;47/Binding site: carbohydrate (Asn) (covalent) #status absent
F;87,168,169,204,280/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
F;197/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F;376-377/Cleavage site: Met-Lys (kallikrein) #status experimental
F;380/Modified site: 4-hydroxyproline (Pro) #status predicted
F;386-387/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;396,400,404,510/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;397,398,518,522,534,546,551,568/Binding site: carbohydrate (Thr) (covalent) #status
F;496-497/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match 68.3%; Score 41; DB 1; Length 619;
Best Local Similarity 72.7%; Pred. No. 9.1;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy      2 LTHTITKLNAE 12
          | |:| |||||
Db    279 LNHSIAKLNAE 289
```

RESULT 8
S05381
VGF8a protein precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 13-Nov-1998
C;Accession: S05381
R;Possenti, R.; Eldridge, J.D.; Paterson, B.M.; Grasso, A.; Levi, A.
EMBO J. 8, 2217-2223, 1989
A;Title: A protein induced by NGF in PC12 cells is stored in secretory vesicles and r
A;Reference number: S05381; MUID:90005425
A;Accession: S05381
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-711 <POS>
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-711/Product: VGF8a protein #status predicted <MAT>

Query Match 68.3%; Score 41; DB 2; Length 711;
Best Local Similarity 58.3%; Pred. No. 11;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 TLTHTITKLNAE 12
          | |||:|::| |
Db    189 TRTHTLTRVNLE 200
```

RESULT 9
F82615
methyltransferase XF1968 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: F82615
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq

Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: F82615
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-534 <SIM>
A;Cross-references: GB:AE004016; GB:AE003849; NID:g9107070; PIDN:AAF84770.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1968

Query Match 65.0%; Score 39; DB 2; Length 534;
Best Local Similarity 58.3%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNAE 12
| | : |||||
Db 359 TTAHAVAKLNAE 370

RESULT 10
S49776
hypothetical protein YDR179w-a - yeast (Saccharomyces cerevisiae) (fragment)
N;Alternate names: hypothetical protein YD9395.13
C;Species: Saccharomyces cerevisiae
C;Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 29-Oct-1999
C;Accession: S49776; S51167
R;Murphy, L.; Harris, D.E.
submitted to the EMBL Data Library, November 1994
A;Reference number: S49764
A;Accession: S49776
A;Molecule type: DNA
A;Residues: 1-498 <MUR>
A;Cross-references: EMBL:Z46727; NID:g1289283; PIDN:CAA86685.1; PID:e223726; PID:g128929
A;Note: the sequence extends from the previous stop codon and does not begin with a star
A;Accession: S51167
A;Molecule type: DNA
A;Residues: 231-498 <MUW>
A;Cross-references: EMBL:Z46727; NID:g1289283; PID:e223644; PID:g1289296; GSPDB:GN00004;
A;Note: the reading frame starts at the first ATG codon
C;Genetics:
A;Gene: MIPS:YDR179w-a
A;Map position: 4R

Query Match 63.3%; Score 38; DB 2; Length 498;
Best Local Similarity 72.7%; Pred. No. 27;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNA 11
|| | |:||||
Db 374 TLQHWISKLNA 384

RESULT 11
A60526
complement C3 - axolotl (fragments)
C;Species: Ambystoma mexicanum (axolotl)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 11-May-2000
C;Accession: A60526
R;Avila, D.; Lambris, J.D.
Comp. Biochem. Physiol. B 95, 839-845, 1990
A;Title: Isolation and characterization of the third complement component of axolotl
A;Reference number: A60526; MUID:90263410
A;Accession: A60526
A;Molecule type: protein
A;Residues: 1-90 <AVI>
C;Superfamily: alpha-2-macroglobulin
C;Keywords: complement pathway; cytolysis; glycoprotein; plasma

Query Match 61.7%; Score 37; DB 2; Length 90;
Best Local Similarity 63.6%; Pred. No. 6.8;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNA 11
|||| :| ||
Db 79 TLTHVVTVYNA 89

RESULT 12
T03013
hypothetical protein 11 - Salmonella typhimurium
C;Species: Salmonella typhimurium
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 08-Oct-1999
C;Accession: T03013
R;Figueroa-Bossi, N.; Bossi, L.
submitted to the EMBL Data Library, June 1998
A;Reference number: Z14818
A;Accession: T03013
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-103 <FIG>
A;Cross-references: EMBL:AF001386; NID:g3294471; PIDN:AAC26074.1; PID:g3294485

Query Match 61.7%; Score 37; DB 2; Length 103;
Best Local Similarity 58.3%; Pred. No. 7.9;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNAE 12
|| ||: | :||
Db 44 TLAHTVEKRDAE 55

RESULT 13
AI0618
conserved hypothetical bacteriophage protein STY1025 [imported] - Salmonella enterica
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C;Accession: AI0618
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
A;Reference number: AB0502; PMID:11677608
A;Accession: AI0618
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-174 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05419.1; PID:g16502180; GSPDB:GN00176
C;Genetics:
A;Gene: STY1025

Query Match 61.7%; Score 37; DB 2; Length 174;

```
  Best Local Similarity   58.3%;  Pred. No. 14;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TLTHTITKLNAE 12
          || ||: | :||
Db     44 TLAHTVEKRDAE 55


RESULT  14
T23883
hypothetical protein R03G8.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T23883
R;Coles, L.
submitted to the EMBL Data Library, February 1996
A;Reference number: Z19813
A;Accession: T23883
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-786 <WIL>
A;Cross-references: EMBL:Z69794; PIDN:CAA93681.1; GSPDB:GN00028; CESP:R03G8.4
A;Experimental source: clone R03G8
C;Genetics:
A;Gene: CESP:R03G8.4
A;Map position: X
A;Introns: 21/2; 115/2; 156/2; 247/3; 309/2; 376/3; 414/3; 585/3; 658/3; 734/2


  Query Match             61.7%;  Score 37;  DB 2;  Length 786;
  Best Local Similarity   63.6%;  Pred. No. 68;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 LTHTITKLNAE 12
          ||| | ||| :
Db    636 LTHAIAKLNQD 646


RESULT  15
S02771
myosin heavy chain A [similarity] - Caenorhabditis elegans
N;Contains: myosin ATPase (EC 3.6.1.32)
C;Species: Caenorhabditis elegans
C;Date: 31-Dec-1993 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C;Accession: T23622; S02771
R;Harris, B.
submitted to the EMBL Data Library, August 1996
A;Reference number: Z19773
A;Accession: T23622
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1992 <WIL>
A;Cross-references: EMBL:Z78199; PIDN:CAB01576.1; GSPDB:GN00023; CESP:K12F2.1
A;Experimental source: clone K12F2
R;Dibb, N.J.; Maruyama, I.N.; Krause, M.; Karn, J.
J. Mol. Biol. 205, 603-613, 1989
A;Title: Sequence analysis of the complete Caenorhabditis elegans myosin heavy chain gen
A;Reference number: S02771; MUID:89178677
A;Accession: S02771
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-116,140-1992 <DIB>
A;Cross-references: EMBL:X08067; NID:g6798; PIDN:CAA30856.1; PID:g6799
C;Genetics:
A;Gene: myo-3; CESP:K12F2.1
A;Map position: 5
A;Introns: 46/1; 192/1; 292/1; 468/2; 1921/3
C;Superfamily: myosin heavy chain; myosin motor domain homology
C;Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; muscle co
F;89-802/Domain: myosin motor domain homology <MMOT>
F;202-209/Region: nucleotide-binding motif A (P-loop)
F;690-712/Region: actin binding #status predicted
F;793-807/Region: actin binding #status predicted
F;875-1992/Domain: coiled coil #status predicted <COI>
F;875-1189/Region: S2
F;1190-1992/Region: light meromyosin
F;153/Modified site: N6,N6,N6-trimethyllysine (Lys) #status predicted
F;208/Binding site: ATP (Lys) #status predicted
F;730,740/Active site: Cys #status predicted


  Query Match             61.7%;  Score 37;  DB 1;  Length 1992;
  Best Local Similarity   72.7%;  Pred. No. 1.8e+02;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 LTHTITKLNAE 12
          ||  |:|||||
Db   1393 LTRQISKLNAE 1403


Search completed: July  1, 2002, 16:20:37
Job time: 201 sec
```

GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:30:12 ; Search time 21.51 Seconds
(without alignments)
21.601 Million cell updates/sec

Title: US-09-461-061A-3
Perfect score: 56
Sequence: 1 IDNVKKARVQVV 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 56 | 100.0 | 644 | 1 | KNG_HUMAN | P01042 homo sapien |
| 2 | 44 | 78.6 | 434 | 1 | KNL2_BOVIN | P01047 bos taurus |
| 3 | 44 | 78.6 | 436 | 1 | KNL1_BOVIN | P01046 bos taurus |
| 4 | 44 | 78.6 | 619 | 1 | KNH2_BOVIN | P01045 bos taurus |
| 5 | 44 | 78.6 | 621 | 1 | KNH1_BOVIN | P01044 bos taurus |
| 6 | 38 | 67.9 | 430 | 1 | KNT1_RAT | P01048 rattus norv |
| 7 | 38 | 67.9 | 430 | 1 | KNT2_RAT | P08932 rattus norv |
| 8 | 38 | 67.9 | 639 | 1 | KNG_RAT | P08934 rattus norv |
| 9 | 38 | 67.9 | 661 | 1 | KNG_MOUSE | O08677 mus musculu |
| 10 | 36 | 64.3 | 322 | 1 | SNF4_YEAST | P12904 saccharomyc |
| 11 | 35 | 62.5 | 367 | 1 | LIVJ_CITFR | P25399 citrobacter |
| 12 | 35 | 62.5 | 551 | 1 | TCPE_YEAST | P40413 saccharomyc |
| 13 | 34 | 60.7 | 203 | 1 | AX2E_PHAAU | O24543 phaseolus a |
| 14 | 34 | 60.7 | 227 | 1 | VATE_SCHPO | O13687 schizosacch |
| 15 | 34 | 60.7 | 249 | 1 | GRPE_SYNY3 | Q59978 synechocyst |
| 16 | 34 | 60.7 | 262 | 1 | YJJV_HAEIN | P44500 haemophilus |
| 17 | 34 | 60.7 | 400 | 1 | AATA_RHIME | Q02635 rhizobium m |
| 18 | 34 | 60.7 | 400 | 1 | AAT_RHILP | O86459 rhizobium l |
| 19 | 34 | 60.7 | 424 | 1 | LE22_METJA | P81291 methanococc |
| 20 | 34 | 60.7 | 566 | 1 | SYR_METJA | Q57689 methanococc |
| 21 | 34 | 60.7 | 679 | 1 | GR78_KLULA | P22010 kluyveromyc |
| 22 | 34 | 60.7 | 1475 | 1 | APU_THETY | P16950 t amylopull |
| 23 | 33.5 | 59.8 | 3164 | 1 | POLG_TUMVJ | P89509 t genome po |
| 24 | 33 | 58.9 | 197 | 1 | GRPE_SYNP7 | Q59984 synechococc |
| 25 | 33 | 58.9 | 208 | 1 | UPP_ECOLI | P25532 escherichia |
| 26 | 33 | 58.9 | 219 | 1 | GCH2_PASMU | P57863 pasteurella |
| 27 | 33 | 58.9 | 306 | 1 | DDG_ECOLI | P76522 escherichia |
| 28 | 33 | 58.9 | 341 | 1 | BIOB_METSK | P94966 methylobaci |
| 29 | 33 | 58.9 | 342 | 1 | BMAP_HUMAN | Q9uk28 homo sapien |
| 30 | 33 | 58.9 | 531 | 1 | SIS2_CANTR | Q12600 candida tro |
| 31 | 33 | 58.9 | 609 | 1 | GLPO_LACLA | Q9cg65 lactococcus |
| 32 | 33 | 58.9 | 663 | 1 | CIRA_ECOLI | P17315 escherichia |
| 33 | 33 | 58.9 | 804 | 1 | GYRB_CHLTR | O84193 chlamydia t |
| 34 | 33 | 58.9 | 1267 | 1 | DHR1_YEAST | Q04217 saccharomyc |
| 35 | 33 | 58.9 | 3305 | 1 | APLP_MANSE | Q25490 manduca sex |
| 36 | 32 | 57.1 | 112 | 1 | GLNB_MYCTU | Q10960 mycobacteri |
| 37 | 32 | 57.1 | 119 | 1 | ACPS_LACLA | Q9ch95 lactococcus |
| 38 | 32 | 57.1 | 191 | 1 | CLD7_RAT | Q9z111 rattus norv |
| 39 | 32 | 57.1 | 211 | 1 | CLD7_HUMAN | O95471 homo sapien |
| 40 | 32 | 57.1 | 219 | 1 | Y413_RICPR | Q9zdb9 rickettsia |
| 41 | 32 | 57.1 | 221 | 1 | HUPD_THIRO | Q56362 thiocapsa r |
| 42 | 32 | 57.1 | 247 | 1 | VATD_BOVIN | P39942 bos taurus |
| 43 | 32 | 57.1 | 247 | 1 | VATD_HUMAN | Q9y5k8 homo sapien |
| 44 | 32 | 57.1 | 247 | 1 | VATD_MOUSE | P57746 mus musculu |
| 45 | 32 | 57.1 | 247 | 1 | VATD_RABIT | O97755 oryctolagus |

ALIGNMENTS

```
RESULT   1
KNG_HUMAN
ID   KNG_HUMAN      STANDARD;      PRT;   644 AA.
AC   P01042; P01043;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Kininogen precursor (Alpha-2-thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
GN   KNG.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC   TISSUE=Liver;
RX   MEDLINE=85234582; PubMed=2989293;
RA   Takagaki Y., Kitamura N., Nakanishi S.;
RT   "Cloning and sequence analysis of cDNAs for human high molecular
RT   weight and low molecular weight prekininogens. Primary structures of
RT   two human prekininogens.";
RL   J. Biol. Chem. 260:8601-8609(1985).
RN   [2]
RP   GENE STRUCTURE.
RX   MEDLINE=85234583; PubMed=2989294;
RA   Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T.,
RA   Nakanishi S.;
RT   "Structural organization of the human kininogen gene and a model for
RT   its evolution.";
RL   J. Biol. Chem. 260:8610-8617(1985).
RN   [3]
RP   SEQUENCE OF 1-401 FROM N.A.
RX   MEDLINE=85122621; PubMed=6441591;
RA   Ohkubo I., Kurachi K., Takasawa T., Shiokawa H., Sasaki M.;
RT   "Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and
RT   its identity with low molecular weight kininogen.";
RL   Biochemistry 23:5691-5697(1984).
RN   [4]
RP   SEQUENCE OF 379-644.
RX   MEDLINE=86030270; PubMed=4054110;
RA   Lottspeich F., Kellermann J., Henschen A., Foertsch B.,
RA   Mueller-Esterl W.;
RT   "The amino acid sequence of the light chain of human high-molecular-
RT   mass kininogen.";
RL   Eur. J. Biochem. 152:307-314(1985).
RN   [5]
RP   SEQUENCE OF 381-389.
RX   MEDLINE=90255622; PubMed=4952632;
RA   Pierce J.V.;
RT   "Structural features of plasma kinins and kininogens.";
RL   Fed. Proc. 27:52-57(1968).
RN   [6]
RP   DISULFIDE BONDS.
RA   Sueyoshi T., Miyata T., Kato H., Iwanaga S.;
RT   "Disulfide bonds in bovine HMW kininogens.";
```

```
RL   Seikagaku 56:808-808(1984).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC       HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
CC       FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-
CC       INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
CC       BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
CC       PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
CC       CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
CC       DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
CC       MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
CC       PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
CC       OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
CC       A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
CC       INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
CC       KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
CC       CLOTTING.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; K02566; AAA35497.1; -.
DR   EMBL; M11437; AAB59550.1; -.
DR   EMBL; M11438; AAB59550.1; JOINED.
DR   EMBL; M11521; AAB59550.1; JOINED.
DR   EMBL; M11522; AAB59550.1; JOINED.
DR   EMBL; M11523; AAB59550.1; JOINED.
DR   EMBL; M11524; AAB59550.1; JOINED.
DR   EMBL; M11525; AAB59550.1; JOINED.
DR   EMBL; M11526; AAB59550.1; JOINED.
DR   EMBL; M11527; AAB59550.1; JOINED.
DR   EMBL; M11528; AAB59550.1; JOINED.
DR   EMBL; M11437; AAB59551.1; -.
DR   EMBL; M11438; AAB59551.1; JOINED.
DR   EMBL; M11521; AAB59551.1; JOINED.
DR   EMBL; M11522; AAB59551.1; JOINED.
DR   EMBL; M11523; AAB59551.1; JOINED.
DR   EMBL; M11524; AAB59551.1; JOINED.
DR   EMBL; M11525; AAB59551.1; JOINED.
DR   EMBL; M11526; AAB59551.1; JOINED.
DR   EMBL; M11527; AAB59551.1; JOINED.
DR   EMBL; M11528; AAB59551.1; JOINED.
DR   PIR; A01279; KGHUH1.
DR   PIR; A25276; A25276.
DR   PIR; A01280; KGHUL1.
DR   PIR; B25276; B25276.
DR   PIR; S02482; S02482.
DR   SWISS-2DPAGE; P01042; HUMAN.
DR   MIM; 228960; -.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW   Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW   Alternative splicing.
FT   SIGNAL       1     18
FT   CHAIN        19    644       KININOGEN.
FT   CHAIN        19    380       KININOGEN HEAVY CHAIN.
FT   PEPTIDE      381   389       BRADYKININ.
FT   CHAIN        390   644       KININOGEN LIGHT CHAIN.
FT   DOMAIN       19    136       CYSTATIN-LIKE 1.
FT   DOMAIN       137   258       CYSTATIN-LIKE 2.
FT   DOMAIN       259   380       CYSTATIN-LIKE 3.
FT   DOMAIN       420   510       HIS-RICH (ASSOCIATED WITH CLOTTING
FT                                ACTIVITY).
FT   REPEAT       420   449
FT   REPEAT       450   479
FT   REPEAT       480   510
FT   MOD_RES      19    19        PYRROLIDONE CARBOXYLIC ACID.
FT   DISULFID     28    614       INTERCHAIN.
FT   DISULFID     83    94
FT   DISULFID     107   126
FT   DISULFID     142   145
FT   DISULFID     206   218
FT   DISULFID     229   248
FT   DISULFID     264   267
FT   DISULFID     328   340
FT   DISULFID     351   370
FT   CARBOHYD     48    48        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     169   169       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     205   205       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     294   294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     401   401       O-LINKED.
FT   CARBOHYD     533   533       O-LINKED.
FT   CARBOHYD     542   542       O-LINKED.
FT   CARBOHYD     546   546       O-LINKED.
FT   CARBOHYD     557   557       O-LINKED.
FT   CARBOHYD     571   571       O-LINKED.
FT   CARBOHYD     577   577       O-LINKED.
FT   CARBOHYD     593   593       O-LINKED.
FT   CARBOHYD     628   628       O-LINKED.
FT   VARSPLIC     402   427       VSPPHTSMAPAQDEERDSGKEQGHTR -> SHLRSCEYKGR
FT                                PPKAGAEPASEREVS (IN ISOFORM LMW).
FT   VARSPLIC     428   644       MISSING (IN ISOFORM LMW).
FT   CONFLICT     593   593       T -> I (IN REF. 1).
SQ   SEQUENCE   644 AA;  71945 MW;  3132B4CBAF8FBB7E CRC64;


  Query Match            100.0%;  Score 56;  DB 1;  Length 644;
  Best Local Similarity  100.0%;  Pred. No. 0.0071;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           ||||||||||||
Db     301 IDNVKKARVQVV 312


RESULT   2
KNL2_BOVIN
ID   KNL2_BOVIN     STANDARD;      PRT;   434 AA.
AC   P01047;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Kininogen, LMW II precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83117859; PubMed=6572010;
RA   Nawa H., Kitamura N., Hirose T., Asai M., Inayama S., Nakanishi S.;
RT   "Primary structures of bovine liver low molecular weight kininogen
RT   precursors and their two mRNAs.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:90-94(1983).
RN   [2]
```

```
RP   SEQUENCE OF 19-376.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (3) THE
CC       ACTIVE PEPTIDE KALLIDIN THAT IS RELEASED FROM LMW-KININOGEN SHOWS
CC       A VARIETY OF PHYSIOLOGICAL EFFECTS: (3A) INFLUENCE IN SMOOTH
CC       MUSCLE CONTRACTION, (3B) INDUCTION OF HYPOTENSION, (3C)
CC       NATRIURESIS AND DIURESIS (KIDNEY).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS: HMW II AND LMW II KININOGEN PRECURSORS ARE
CC       PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
CC       SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
CC       TO RESIDUE 398.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- MISCELLANEOUS: LMW-KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT
CC       INVOLVED IN BLOOD CLOTTING.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00427; CAA23710.1; -.
DR   PIR; A01284; KGBOL2.
DR   HSSP; P01038; 1A90.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   Pfam; PF00031; cystatin; 3.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Signal.
FT   SIGNAL        1     18
FT   CHAIN        19    434       KININOGEN, LMW II.
FT   CHAIN        19    376       HEAVY CHAIN.
FT   PEPTIDE     378    386       BRADYKININ.
FT   CHAIN       387    434       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    256       CYSTATIN-LIKE 2.
FT   DOMAIN      257    376       CYSTATIN-LIKE 3.
FT   MOD_RES      19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    136    136       O-LINKED (PARTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .); OR 169.
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    280    280       N-LINKED (GLCNAC. . .).
FT   DISULFID     27    404       INTERCHAIN.
FT   DISULFID     82     93
FT   DISULFID    106    125
FT   DISULFID    141    144
FT   DISULFID    205    217
FT   DISULFID    228    247
FT   DISULFID    261    264
FT   DISULFID    325    337
FT   DISULFID    348    367
SQ   SEQUENCE   434 AA;  48148 MW;  73A7079DE3E03430 CRC64;
```

```
  Query Match             78.6%;  Score 44;  DB 1;  Length 434;
  Best Local Similarity   83.3%;  Pred. No. 0.8;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           || |||| ||||
Db     298 IDTVKKATVQVV 309
```

```
RESULT   3
KNL1_BOVIN
ID   KNL1_BOVIN     STANDARD;      PRT;   436 AA.
AC   P01046;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Kininogen, LMW I precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83117859; PubMed=6572010;
RA   Nawa H., Kitamura N., Hirose T., Asai M., Inayama S., Nakanishi S.;
RT   "Primary structures of bovine liver low molecular weight kininogen
RT   precursors and their two mRNAs.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:90-94(1983).
RN   [2]
RP   SEQUENCE OF 19-378.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (3) THE
CC       ACTIVE PEPTIDE KALLIDIN THAT IS RELEASED FROM LMW-KININOGEN SHOWS
CC       A VARIETY OF PHYSIOLOGICAL EFFECTS: (3A) INFLUENCE IN SMOOTH
CC       MUSCLE CONTRACTION, (3B) INDUCTION OF HYPOTENSION, (3C)
CC       NATRIURESIS AND DIURESIS (KIDNEY).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS: HMW I AND LMW I KININOGEN PRECURSORS ARE
CC       PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
CC       SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
CC       TO RESIDUE 400.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- MISCELLANEOUS: LMW-KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT
CC       INVOLVED IN BLOOD CLOTTING.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00426; CAA23709.1; -.
DR   EMBL; J00010; AAA30604.1; -.
DR   PIR; A01283; KGBOL1.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   Pfam; PF00031; cystatin; 3.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Signal.
```

```
FT   SIGNAL        1     18
FT   CHAIN         19    436       KININOGEN, LMW I.
FT   CHAIN         19    378       HEAVY CHAIN.
FT   PEPTIDE      380    388       BRADYKININ.
FT   CHAIN        389    436       LIGHT CHAIN.
FT   DOMAIN        19    135       CYSTATIN-LIKE 1.
FT   DOMAIN       136    257       CYSTATIN-LIKE 2.
FT   DOMAIN       258    378       CYSTATIN-LIKE 3.
FT   MOD_RES       19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD      87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD     136    136       O-LINKED (PARTIAL).
FT   CARBOHYD     168    168       N-LINKED (GLCNAC. . .); OR 169.
FT   CARBOHYD     197    197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD     204    204       N-LINKED (GLCNAC. . .).
FT   DISULFID      27    406       INTERCHAIN.
FT   DISULFID      82     93
FT   DISULFID     106    125
FT   DISULFID     141    144
FT   DISULFID     205    217
FT   DISULFID     228    247
FT   DISULFID     263    266
FT   DISULFID     327    339
FT   DISULFID     350    369
FT   CONFLICT     295    295       A -> T (IN REF. 1; CAA23709).
SQ   SEQUENCE   436 AA;  48427 MW;  F01F7EB6814BCE6C CRC64;


  Query Match          78.6%;  Score 44;  DB 1;  Length 436;
  Best Local Similarity   83.3%;  Pred. No. 0.8;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARVQVV 12
          || |||| ||||
Db    300 IDTVKKATVQVV 311


RESULT   4
KNH2_BOVIN
ID   KNH2_BOVIN     STANDARD;      PRT;   619 AA.
AC   P01045;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Kininogen, HMW II precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84014106; PubMed=6571699;
RA   Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
RT   "A single gene for bovine high molecular weight and low molecular
RT   weight kininogens.";
RL   Nature 305:545-549(1983).
RN   [2]
RP   SEQUENCE OF 19-376.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
RN   [3]
RP   SEQUENCE OF 376-391.
RX   MEDLINE=70180420; PubMed=4986212;
RA   Kato H., Nagasawa S., Suzuki T.;
RT   "Studies on the structure of bovine kininogen: cleavages of disulfide
RT   bonds and of methionyl bonds in kininogen-II.";
RL   J. Biochem. 67:313-323(1970).
RN   [4]
RP   SEQUENCE OF 387-455.
RX   MEDLINE=76260155; PubMed=956151;
RA   Han Y.N., Kato H., Iwanaga S., Suzuki T.;
RT   "Primary structure of bovine plasma high-molecular-weight kininogen.
RT   The amino acid sequence of a glycopeptide portion (fragment 1)
RT   following the C-terminus ot the bradykinin moiety.";
RL   J. Biochem. 79:1201-1222(1976).
RN   [5]
RP   SEQUENCE OF 456-496.
RX   MEDLINE=75170265; PubMed=1169237;
RA   Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
RT   "Studies on the primary structure of bovine high-molecular-weight
RT   kininogen. Amino acid sequence of a fragment ('histidine-rich
RT   peptide') released by plasma kallikrein.";
RL   J. Biochem. 77:55-68(1975).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC       HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
CC       TO FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
CC       PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE
CC       PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
CC       A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
CC       MUSCLE CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C)
CC       NATRIURESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
CC       (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
CC       IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
CC       RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
CC       (4F) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
CC       ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
CC       ACTION).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS: HMW II AND LMW II KININOGEN PRECURSORS ARE
CC       PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
CC       SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
CC       TO RESIDUE 398.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; V01492; CAA24736.1; -.
DR   EMBL; V01492; CAA24737.1; ALT_SEQ.
DR   PIR; A01282; KGBOH2.
DR   PIR; B29559; B29559.
DR   HSSP; P01038; 1A90.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Blood coagulation; Signal;
KW   Inflammatory response.
FT   SIGNAL        1     18
FT   CHAIN         19    619       KININOGEN, HMW II.
FT   CHAIN         19    376       HEAVY CHAIN.
FT   PEPTIDE      378    386       BRADYKININ.
FT   CHAIN        387    619       LIGHT CHAIN.
FT   DOMAIN        19    135       CYSTATIN-LIKE 1.
FT   DOMAIN       136    256       CYSTATIN-LIKE 2.
FT   DOMAIN       257    376       CYSTATIN-LIKE 3.
FT   MOD_RES       19     19       PYRROLIDONE CARBOXYLIC ACID.
```

```
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    136    136       O-LINKED (PARTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .); OR 169.
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    280    280       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    400    400       O-LINKED.
FT   DISULFID     27    589       INTERCHAIN.
FT   DISULFID     82     93
FT   DISULFID    106    125
FT   DISULFID    141    144
FT   DISULFID    205    217
FT   DISULFID    228    247
FT   DISULFID    261    264
FT   DISULFID    325    337
FT   DISULFID    348    367
FT   VARIANT     398    398       T -> P.
FT   VARIANT     401    401       L -> V.
FT   VARIANT     454    454       H -> K.
SQ   SEQUENCE   619 AA;  68710 MW;  F04320A8EB0EE0DA CRC64;


  Query Match             78.6%;  Score 44;  DB 1;  Length 619;
  Best Local Similarity   83.3%;  Pred. No. 1.1;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           || |||| ||||
Db     298 IDTVKKATVQVV 309


RESULT   5
KNH1_BOVIN
ID   KNH1_BOVIN     STANDARD;      PRT;   621 AA.
AC   P01044;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Kininogen, HMW I precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84014106; PubMed=6571699;
RA   Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
RT   "A single gene for bovine high molecular weight and low molecular
RT   weight kininogens.";
RL   Nature 305:545-549(1983).
RN   [2]
RP   SEQUENCE OF 19-378.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
RN   [3]
RP   SEQUENCE OF 378-393.
RX   MEDLINE=70180420; PubMed=4986212;
RA   Kato H., Nagasawa S., Suzuki T.;
RT   "Studies on the structure of bovine kininogen: cleavages of disulfide
RT   bonds and of methionyl bonds in kininogen-II.";
RL   J. Biochem. 67:313-323(1970).
RN   [4]
RP   SEQUENCE OF 458-498.
RX   MEDLINE=75170265; PubMed=1169237;
RA   Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
RT   "Studies on the primary structure of bovine high-molecular-weight
RT   kininogen. Amino acid sequence of a fragment ('histidine-rich
RT   peptide') released by plasma kallikrein.";
RL   J. Biochem. 77:55-68(1975).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC       HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
CC       TO FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
CC       PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE
CC       PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
CC       A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
CC       MUSCLE CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C)
CC       NATRIURESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
CC       (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
CC       IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
CC       RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
CC       (4F) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
CC       ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
CC       ACTION).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS: HMW I AND LMW I KININOGEN PRECURSORS ARE
CC       PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
CC       SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
CC       TO RESIDUE 400.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; V01491; CAA24735.1; -.
DR   PIR; A01281; KGBOH1.
DR   PIR; A29559; A29559.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Blood coagulation;
KW   Inflammatory response; Signal.
FT   SIGNAL       1     18       PROBABLE.
FT   CHAIN       19    621       KININOGEN, HMW I.
FT   CHAIN       19    378       HEAVY CHAIN.
FT   PEPTIDE    380    388       BRADYKININ.
FT   CHAIN      389    621       LIGHT CHAIN.
FT   DOMAIN      19    135       CYSTATIN-LIKE 1.
FT   DOMAIN     136    257       CYSTATIN-LIKE 2.
FT   DOMAIN     258    378       CYSTATIN-LIKE 3.
FT   MOD_RES     19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD    87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   136    136       O-LINKED (PARTIAL).
FT   CARBOHYD   168    168       N-LINKED (GLCNAC. . .); OR 169.
FT   CARBOHYD   197    197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD   204    204       N-LINKED (GLCNAC. . .).
FT   DISULFID    27    591       INTERCHAIN.
FT   DISULFID    82     93
FT   DISULFID   106    125
FT   DISULFID   141    144
FT   DISULFID   205    217
FT   DISULFID   228    247
FT   DISULFID   263    266
FT   DISULFID   327    339
FT   DISULFID   350    369
SQ   SEQUENCE   621 AA;  68890 MW;  D16850BEFE3C55CD CRC64;
```

```
  Query Match             78.6%;  Score 44;  DB 1;  Length 621;
  Best Local Similarity   83.3%;  Pred. No. 1.1;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           || |||| ||||
Db     300 IDTVKKATVQVV 311


RESULT   6
KNT1_RAT
ID   KNT1_RAT       STANDARD;      PRT;   430 AA.
AC   P01048; P04081;
DT   01-NOV-1986 (Rel. 03, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   T-kininogen I precursor (Major acute phase protein) (Alpha-1-MAP)
DE   (Thiostatin) [Contains: T-kinin].
GN   MAP1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86008264; PubMed=2413018;
RA   Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT   "Primary structures of the mRNAs encoding the rat precursors for
RT   bradykinin and T-kinin. Structural relationship of kininogens with
RT   major acute phase protein and alpha 1-cysteine proteinase
RT   inhibitor.";
RL   J. Biol. Chem. 260:12054-12059(1985).
RN   [2]
RP   SEQUENCE OF 5-430 FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=86008266; PubMed=2413019;
RA   Anderson K.P., Heath E.C.;
RT   "The relationship between rat major acute phase protein and the
RT   kininogens.";
RL   J. Biol. Chem. 260:12065-12071(1985).
RN   [3]
RP   SEQUENCE OF 7-430 FROM N.A.
RX   MEDLINE=85127561; PubMed=2578992;
RA   Cole T., Inglis A.S., Roxburgh C.M., Howlett G.J., Schreiber G.;
RT   "Major acute phase alpha 1-protein of the rat is homologous to bovine
RT   kininogen and contains the sequence for bradykinin: its synthesis is
RT   regulated at the mRNA level.";
RL   FEBS Lett. 182:57-61(1985).
RN   [4]
RP   SEQUENCE OF 1-65 FROM N.A.
RX   MEDLINE=87250580; PubMed=2439509;
RA   Fung W.-P., Schreiber G.;
RT   "Structure and expression of the genes for major acute phase alpha 1-
RT   protein (thiostatin) and kininogen in the rat.";
RL   J. Biol. Chem. 262:9298-9308(1987).
CC   -!- FUNCTION: KININOGENS ARE PLASMA GLYCOPROTEINS WITH A NUMBER OF
CC       FUNCTIONS: (1) AS PRECURSOR OF THE ACTIVE PEPTIDE BRADYKININ THEY
CC       EFFECT SMOOTH MUSCLE CONTRACTION, INDUCTION OF HYPOTENSION AND
CC       INCREASE OF VASCULAR PERMEABILITY. (2) THEY PLAY A ROLE IN BLOOD
CC       COAGULATION BY HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND
CC       FACTOR XI NEXT TO FACTOR XII. (3) THEY ARE INHIBITOR OF THIOL
CC       PROTEASES.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- INDUCTION: IN RESPONSE TO AN INFLAMMATORY STIMULANT. T-KININOGEN
CC       II SYNTHESIS IS INDUCED AND THE PLASMA CONCENTRATION OF
CC       T-KININOGEN I IS RAISED.
CC   -!- PTM: AS T-KININ IS PRECEDED BY A MET INSTEAD OF AN ARG OR LYS, IT
CC       IS NOT RELEASED FROM ITS PRECURSOR BY EITHER TISSUE OR PLASMA
CC       KALLIKREIN.
CC   -!- MISCELLANEOUS: RAT EXPRESS FOUR TYPES OF KININOGENS: THE CLASSICAL
CC       HMW AND LMW KININOGENS PRODUCED BY ALTERNATIVE SPLICING OF THE
CC       SAME GENE, AND TWO ADDITIONAL LMW-LIKE KININOGENS: T-I AND T-II.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   -!- CAUTION: IN ADDITION TO THE CONFLICTS DESCRIBED IN THE FEATURE
CC       TABLE, REF.2. SEQUENCE DIFFERS FROM THAT SHOWN IN POSITIONS: 257,
CC       262,268,269,295,314,315,331,332, AND 389. IN ALL THOSE POSITIONS
CC       THE ALTERNATE AMINO-ACID IS THE ONE PRESENT IN T-II KININOGEN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M11883; AAA41489.1; -.
DR   EMBL; M11661; AAA41570.1; -.
DR   EMBL; M16454; AAA41568.1; -.
DR   EMBL; X02299; CAA26162.1; ALT_SEQ.
DR   PIR; A01286; KGRTT1.
DR   PIR; A01285; KGRTM.
DR   PIR; A23897; A23897.
DR   PIR; A27115; A27115.
DR   GlycoSuiteDB; P01048; -.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   Pfam; PF00031; cystatin; 3.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Multigene family;
KW   Thiol protease inhibitor; Bradykinin; Acute phase; Signal.
FT   SIGNAL        1     18
FT   CHAIN        19    430       KININOGEN, T-I.
FT   CHAIN        19    375       HEAVY CHAIN.
FT   PEPTIDE     376    386       T-KININ.
FT   CHAIN       387    430       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    257       CYSTATIN-LIKE 2.
FT   DOMAIN      258    375       CYSTATIN-LIKE 3.
FT   DISULFID     28    404       INTERCHAIN (BY SIMILARITY).
FT   DISULFID     83     94       BY SIMILARITY.
FT   DISULFID    107    125       BY SIMILARITY.
FT   DISULFID    141    144       BY SIMILARITY.
FT   DISULFID    205    217       BY SIMILARITY.
FT   DISULFID    228    247       BY SIMILARITY.
FT   DISULFID    263    266       BY SIMILARITY.
FT   DISULFID    327    339       BY SIMILARITY.
FT   DISULFID    350    369       BY SIMILARITY.
FT   CARBOHYD     82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    326    326       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     26     28       LNC -> MDR (IN REF. 2).
FT   CONFLICT     55     55       V -> L (IN REF. 2).
FT   CONFLICT     61     61       E -> K (IN REF. 1).
FT   CONFLICT     83     83       C -> Y (IN REF. 3).
FT   CONFLICT    166    166       S -> F (IN REF. 2 AND 3).
FT   CONFLICT    179    181       REV -> TKI (IN REF. 2).
FT   CONFLICT    193    193       N -> D (IN REF. 2).
FT   CONFLICT    212    212       S -> F (IN REF. 2).
FT   CONFLICT    214    214       R -> H (IN REF. 3).
FT   CONFLICT    229    229       T -> R (IN REF. 2).
FT   CONFLICT    233    233       H -> Y (IN REF. 2).
FT   CONFLICT    257    257       E -> S (IN REF. 2).
FT   CONFLICT    262    262       N -> K (IN REF. 2).
FT   CONFLICT    264    264       R -> F (IN REF. 2).
FT   CONFLICT    268    269       RE -> KN (IN REF. 2).
FT   CONFLICT    295    295       I -> L (IN REF. 2).
FT   CONFLICT    314    315       VI -> TK (IN REF. 2).
FT   CONFLICT    331    332       SK -> TN (IN REF. 2).
FT   CONFLICT    389    389       R -> Q (IN REF. 2).
```

```
FT   CONFLICT    414    414       R -> G (IN REF. 2 AND 3).
FT   CONFLICT    415    415       A -> L (IN REF. 2).
FT   CONFLICT    420    421       DH -> ER (IN REF. 3).
FT   CONFLICT    430    430       P -> S (IN REF. 1).
SQ   SEQUENCE   430 AA;  47715 MW;  FAEBB78FAF4723C3 CRC64;


  Query Match             67.9%;  Score 38;  DB 1;  Length 430;
  Best Local Similarity   75.0%;  Pred. No. 10;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           || ||||  |||
Db     300 IDTVKKATSQVV 311


RESULT   7
KNT2_RAT
ID   KNT2_RAT       STANDARD;      PRT;   430 AA.
AC   P08932;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   T-kininogen II precursor (Major acute phase protein) (Alpha-1-MAP)
DE   (Thiostatin) [Contains: T-kinin].
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86008264; PubMed=2413018;
RA   Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT   "Primary structures of the mRNAs encoding the rat precursors for
RT   bradykinin and T-kinin. Structural relationship of kininogens with
RT   major acute phase protein and alpha 1-cysteine proteinase
RT   inhibitor.";
RL   J. Biol. Chem. 260:12054-12059(1985).
CC   -!- FUNCTION: KININOGENS ARE PLASMA GLYCOPROTEINS WITH A NUMBER OF
CC       FUNCTIONS: (1) AS PRECURSOR OF THE ACTIVE PEPTIDE BRADYKININ THEY
CC       EFFECT SMOOTH MUSCLE CONTRACTION, INDUCTION OF HYPOTENSION AND
CC       INCREASE OF VASCULAR PERMEABILITY. (2) THEY PLAY A ROLE IN BLOOD
CC       COAGULATION BY HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND
CC       FACTOR XI NEXT TO FACTOR XII. (3) THEY ARE INHIBITOR OF THIOL
CC       PROTEASES.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- INDUCTION: IN RESPONSE TO AN INFLAMMATORY STIMULANT. T-KININOGEN
CC       II SYNTHESIS IS INDUCED AND THE PLASMA CONCENTRATION OF
CC       T-KININOGEN I IS RAISED.
CC   -!- PTM: AS T-KININ IS PRECEDED BY A MET INSTEAD OF AN ARG OR LYS, IT
CC       IS NOT RELEASED FROM ITS PRECURSOR BY EITHER TISSUE OR PLASMA
CC       KALLIKREIN.
CC   -!- MISCELLANEOUS: RAT EXPRESS FOUR TYPES OF KININOGENS: THE CLASSICAL
CC       HMW AND LMW KININOGENS PRODUCED BY ALTERNATIVE SPLICING OF THE
CC       SAME GENE, AND TWO ADDITIONAL LMW-LIKE KININOGENS: T-I AND T-II.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M11885; AAA41491.1; -.
DR   PIR; B28055; B28055.
DR   GlycoSuiteDB; P08932; -.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   Pfam; PF00031; cystatin; 3.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Multigene family;
KW   Thiol protease inhibitor; Bradykinin; Acute phase; Signal.
FT   SIGNAL        1     18
FT   CHAIN        19    430       KININOGEN, T-II.
FT   CHAIN        19    375       HEAVY CHAIN.
FT   PEPTIDE     376    386       T-KININ.
FT   CHAIN       387    430       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    257       CYSTATIN-LIKE 2.
FT   DOMAIN      258    375       CYSTATIN-LIKE 3.
FT   DISULFID     28    404       INTERCHAIN (BY SIMILARITY).
FT   DISULFID     83     94       BY SIMILARITY.
FT   DISULFID    107    125       BY SIMILARITY.
FT   DISULFID    141    144       BY SIMILARITY.
FT   DISULFID    205    217       BY SIMILARITY.
FT   DISULFID    228    247       BY SIMILARITY.
FT   DISULFID    263    266       BY SIMILARITY.
FT   DISULFID    327    339       BY SIMILARITY.
FT   DISULFID    350    369       BY SIMILARITY.
FT   CARBOHYD     82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    326    326       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   430 AA;  47524 MW;  43EDF02D1BF55076 CRC64;


  Query Match             67.9%;  Score 38;  DB 1;  Length 430;
  Best Local Similarity   75.0%;  Pred. No. 10;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           || ||||  |||
Db     300 IDTVKKATSQVV 311


RESULT   8
KNG_RAT
ID   KNG_RAT        STANDARD;      PRT;   639 AA.
AC   P08934; P08933;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Kininogen precursor [Contains: Bradykinin].
GN   KNG.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RX   MEDLINE=87137443; PubMed=3029068;
RA   Kitagawa H., Kitamura N., Hayashida H., Miyata T., Nakanishi S.;
RT   "Differing expression patterns and evolution of the rat kininogen
RT   gene family.";
RL   J. Biol. Chem. 262:2190-2198(1987).
RN   [2]
RP   SEQUENCE FROM N.A. (LMW ISOFORM).
RX   MEDLINE=86008264; PubMed=2413018;
RA   Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT   "Primary structures of the mRNAs encoding the rat precursors for
RT   bradykinin and T-kinin. Structural relationship of kininogens with
RT   major acute phase protein and alpha 1-cysteine proteinase
RT   inhibitor.";
RL   J. Biol. Chem. 260:12054-12059(1985).
RN   [3]
RP   SEQUENCE OF 1-65 FROM N.A.
RC   STRAIN=BUFFALO;
RX   MEDLINE=87250580; PubMed=2439509;
RA   Fung W.-P., Schreiber G.;
RT   "Structure and expression of the genes for major acute phase alpha 1-
```

```
RT   protein (thiostatin) and kininogen in the rat.";
RL   J. Biol. Chem. 262:9298-9308(1987).
RN   [4]
RP   SEQUENCE OF 1-41 FROM N.A.
RC   STRAIN=WISTAR; TISSUE=Liver;
RX   MEDLINE=87137465; PubMed=3818598;
RA   Kageyama R., Kitamura N., Ohkubo H., Nakanishi S.;
RT   "Differing utilization of homologous transcription initiation sites
RT   of rat K and T kininogen genes under inflammation condition.";
RL   J. Biol. Chem. 262:2345-2351(1987).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC       HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
CC       FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-
CC       INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
CC       BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
CC       PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
CC       CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
CC       DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
CC       MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
CC       PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
CC       OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
CC       A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
CC       INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
CC       KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
CC       CLOTTING.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- MISCELLANEOUS: RAT EXPRESS FOUR TYPES OF KININOGENS: THE CLASSICAL
CC       HMW/LMW KININOGENS AND TWO ADDITIONAL LMW-LIKE KININOGENS: T-I AND
CC       T-II.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L29428; AAA41486.1; -.
DR   EMBL; M11884; AAA41487.1; -.
DR   EMBL; M14369; AAA41484.1; -.
DR   EMBL; M14369; AAA41485.1; ALT_SEQ.
DR   EMBL; M16455; AAA41482.1; -.
DR   PIR; A25486; A25486.
DR   PIR; A28055; A28055.
DR   HSSP; P01040; 1DVD.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW   Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW   Alternative splicing; Multigene family.
FT   SIGNAL       1     18
FT   CHAIN       19    639       KININOGEN.
FT   CHAIN       19    380       KININOGEN HEAVY CHAIN.
FT   PEPTIDE    381    389       BRADYKININ.
FT   CHAIN      390    639       KININOGEN LIGHT CHAIN.
FT   DOMAIN      19    136       CYSTATIN-LIKE 1.
FT   DOMAIN     137    258       CYSTATIN-LIKE 2.
FT   DOMAIN     259    380       CYSTATIN-LIKE 3.
FT   DOMAIN     439    514       HIS-RICH.
FT   DISULFID    28    609       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    83     94       BY SIMILARITY.
FT   DISULFID   107    126       BY SIMILARITY.
FT   DISULFID   142    145       BY SIMILARITY.
FT   DISULFID   206    218       BY SIMILARITY.
FT   DISULFID   229    248       BY SIMILARITY.
FT   DISULFID   264    267       BY SIMILARITY.
FT   DISULFID   328    340       BY SIMILARITY.
FT   DISULFID   351    370       BY SIMILARITY.
FT   CARBOHYD    82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   127    127       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   169    169       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   205    205       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   294    294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   529    529       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   402    433       VSPSYIARVQEERDPGNEQGPIHGHGWLHAKQ -> RLLNS
FT                               CEYKGRLLKAGAGPAPERQAEASTVTP (IN ISOFORM
FT                               LMW).
FT   VARSPLIC   434    639       MISSING (IN ISOFORM LMW).
FT   CONFLICT    61     61       E -> K (IN REF. 2).
SQ   SEQUENCE   639 AA;  70933 MW;  D3172DF94FF56AF5 CRC64;


  Query Match             67.9%;  Score 38;  DB 1;  Length 639;
  Best Local Similarity   75.0%;  Pred. No. 15;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           || ||||  |||
Db     301 IDTVKKATSQVV 312


RESULT   9
KNG_MOUSE
ID   KNG_MOUSE      STANDARD;      PRT;   661 AA.
AC   O08677; O08676;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Kininogen precursor [Contains: Bradykinin].
GN   KNG.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC   STRAIN=C57BL/6 x CBA; TISSUE=Liver;
RA   Takano M., Kondoh J., Yayama K., Okamoto H.;
RT   "Molecular cloning of cDNAs for mouse low- and high- molecular
RT   kininogen.";
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC       HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
CC       FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-
CC       INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
CC       BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
CC       PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
CC       CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
CC       DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
CC       MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
CC       PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
CC       OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
CC       A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
CC       INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
CC       KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
CC       CLOTTING (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: PLASMA.
```

```
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D84435; BAA19743.1; -.
DR   EMBL; D84415; BAA19742.1; -.
DR   MGD; MGI:1097705; Kng.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 1.
KW   Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW   Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW   Alternative splicing.
FT   SIGNAL       1     18       POTENTIAL.
FT   CHAIN       19    661       KININOGEN.
FT   CHAIN       19    379       KININOGEN HEAVY CHAIN.
FT   CHAIN      380    388       BRADYKININ.
FT   CHAIN      389    661       KININOGEN LIGHT CHAIN.
FT   DOMAIN      19    135       CYSTATIN-LIKE 1.
FT   DOMAIN     136    257       CYSTATIN-LIKE 2.
FT   DOMAIN     258    379       CYSTATIN-LIKE 3.
FT   DOMAIN     439    524       HIS-RICH.
FT   DISULFID    28    631       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    83     94       BY SIMILARITY.
FT   DISULFID   107    125       BY SIMILARITY.
FT   DISULFID   141    144       BY SIMILARITY.
FT   DISULFID   205    217       BY SIMILARITY.
FT   DISULFID   228    247       BY SIMILARITY.
FT   DISULFID   263    266       BY SIMILARITY.
FT   DISULFID   327    339       BY SIMILARITY.
FT   DISULFID   350    369       BY SIMILARITY.
FT   CARBOHYD    82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   168    168       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   204    204       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   242    242       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   401    432       VSPPYIAREQEERDAETEQGPTHGHGWLHEKQ -> RLLRA
FT                               CEYKGRLSKAGAEPAPERQAESSQVKQ (IN ISOFORM
FT                               LMW).
FT   VARSPLIC   433    661       MISSING (IN ISOFORM LMW).
SQ   SEQUENCE   661 AA;  73102 MW;  774460258D58796E CRC64;


  Query Match             67.9%;  Score 38;  DB 1;  Length 661;
  Best Local Similarity   75.0%;  Pred. No. 16;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           || ||||  |||
Db     300 IDTVKKATSQVV 311


RESULT  10
SNF4_YEAST
ID   SNF4_YEAST     STANDARD;      PRT;   322 AA.
AC   P12904;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Nuclear protein SNF4 (Regulatory protein CAT3).
GN   SNF4 OR CAT3 OR YGL115W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89006284; PubMed=3049255;
RA   Schueller H.-J., Entian K.-D.;
RT   "Molecular characterization of yeast regulatory gene CAT3 necessary
RT   for glucose derepression and nuclear localization of its product.";
RL   Gene 67:247-257(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90097921; PubMed=2481228;
RA   Celenza J.L., Eng F.J., Carlson M.;
RT   "Molecular analysis of the SNF4 gene of Saccharomyces cerevisiae:
RT   evidence for physical association of the SNF4 protein with the SNF1
RT   protein kinase.";
RL   Mol. Cell. Biol. 9:5045-5054(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Lauquin G.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 1-21 FROM N.A.
RA   Doi A., Doi K.;
RL   Submitted (JUN-1993) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 30-34 AND 316-322.
RX   MEDLINE=94131988; PubMed=7905477;
RA   Mitchelhill K.I., Stapleton D., Gao G., House C., Michell B.,
RA   Katsis F., Witters L.A., Kemp B.E.;
RT   "Mammalian AMP-activated protein kinase shares structural and
RT   functional homology with the catalytic domain of yeast Snf1 protein
RT   kinase.";
RL   J. Biol. Chem. 269:2361-2364(1994).
CC   -!- FUNCTION: THIS PROTEIN CAUSES EXPRESSION OF GLUCOSE-REPRESSIBLE
CC       GENES UPON GLUCOSE DEPRIVATION. IT INTERACTS AND HAS FUNCTIONAL
CC       RELATIONSHIP TO THE PROTEIN-KINASE SNF1.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA
CC       SUBUNIT FAMILY.
CC   -!- SIMILARITY: CONTAINS 4 CBS DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21760; AAA34472.1; -.
DR   EMBL; M30470; AAA35061.1; -.
DR   EMBL; Z72637; CAA96823.1; -.
DR   EMBL; D16506; BAA03958.1; -.
DR   PIR; JT0316; RGBYC3.
DR   PIR; A33480; A33480.
DR   SGD; S0003083; SNF4.
DR   InterPro; IPR000644; CBS.
DR   Pfam; PF00571; CBS; 4.
DR   SMART; SM00116; CBS; 4.
KW   Carbohydrate metabolism; Transcription regulation; Nuclear protein;
KW   Repeat; CBS domain.
FT   DOMAIN       35     89       CBS 1.
FT   DOMAIN      117    175       CBS 2.
FT   DOMAIN      192    246       CBS 3.
FT   DOMAIN      259    318       CBS 4.
SQ   SEQUENCE   322 AA;  36401 MW;  51B387E346EE9561 CRC64;


  Query Match             64.3%;  Score 36;  DB 1;  Length 322;
  Best Local Similarity   66.7%;  Pred. No. 18;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 IDNVKKARV 9
            :||::||||
Db    284 MDNIRKARV 292


RESULT  11
LIVJ_CITFR
ID   LIVJ_CITFR     STANDARD;      PRT;   367 AA.
AC   P25399;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Leu/Ile/Val-binding protein precursor (LIV-BP).
GN   LIVJ.
OS   Citrobacter freundii.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Citrobacter.
OX   NCBI_TaxID=546;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 8090;
RA   Daggett Garvin L., Hardies S.C.;
RL   Submitted (APR-1991) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE LEUCINE, ISOLEUCINE,
CC       VALINE, (THREONINE) TRANSPORT SYSTEM, WHICH IS ONE OF THE TWO
CC       PERIPLASMIC BINDING PROTEIN-DEPENDENT TRANSPORT SYSTEMS OF THE
CC       HIGH-AFFINITY TRANSPORT OF THE BRANCHED-CHAIN AMINO ACIDS.
CC   -!- SUBCELLULAR LOCATION: Periplasmic.
CC   -!- SIMILARITY: BELONGS TO THE LEUCINE-BINDING PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X58820; CAA41622.1; -.
DR   PIR; S14619; S14619.
DR   HSSP; P02917; 2LIV.
DR   InterPro; IPR001828; ANF_receptor.
DR   InterPro; IPR000709; Leu_Ile_Val_bind.
DR   Pfam; PF01094; ANF_receptor; 2.
DR   PRINTS; PR00337; LEUILEVALBP.
KW   Amino-acid transport; Transport; Periplasmic; Signal.
FT   SIGNAL       1     23       BY SIMILARITY.
FT   CHAIN       24    367       LEU/ILE/VAL-BINDING PROTEIN.
FT   DISULFID    76    101       BY SIMILARITY.
SQ   SEQUENCE   367 AA;  39087 MW;  F731F36CB64CAC19 CRC64;


  Query Match             62.5%;  Score 35;  DB 1;  Length 367;
  Best Local Similarity   63.6%;  Pred. No. 31;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 DNVKKARVQVV 12
          ||:|||   ||
Db    183 DNLKKANADVV 193


RESULT  12
TCPE_YEAST
ID   TCPE_YEAST     STANDARD;      PRT;   551 AA.
AC   P40413;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   T-complex protein 1, epsilon subunit (TCP-1-epsilon) (CCT-epsilon).
GN   CCT5 OR TCP5 OR YJR064W OR J1752.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YPH501;
RA   Kim S.;
RL   Submitted (XXX-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE=96437976; PubMed=8840504;
RA   Huang M.-E., Manus V., Chuat J.-C., Galibert F.;
RT   "Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open
RT   reading frames and a gene cluster with a counterpart on chromosome
RT   XI.";
RL   Yeast 12:869-875(1996).
CC   -!- FUNCTION: MOLECULAR CHAPERONE; ASSIST THE FOLDING OF PROTEINS UPON
CC       ATP HYDROLYSIS. KNOWN TO PLAY A ROLE, IN VITRO, IN THE FOLDING OF
CC       ACTIN AND TUBULIN. IN YEAST MAY PLAY A ROLE IN MITOTIC SPINDLE
CC       FORMATION.
CC   -!- SUBUNIT: HETERO-OLIGOMERIC COMPLEX OF ABOUT 850 TO 900 kDa THAT
CC       FORMS TWO STACKED RINGS, 12 TO 16 NM IN DIAMETER.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE TCP-1 CHAPERONIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L37350; AAA53132.1; -.
DR   EMBL; Z49564; CAA89592.1; ALT_INIT.
DR   EMBL; L47993; AAB39290.1; ALT_INIT.
DR   HSSP; P48425; 1A6D.
DR   SGD; S0003825; CCT5.
DR   InterPro; IPR002423; TCP1_cpn60.
DR   InterPro; IPR002194; TCP_1.
DR   Pfam; PF00118; cpn60_TCP1; 1.
DR   PRINTS; PR00304; TCOMPLEXTCP1.
DR   PROSITE; PS00750; TCP1_1; 1.
DR   PROSITE; PS00751; TCP1_2; FALSE_NEG.
DR   PROSITE; PS00995; TCP1_3; 1.
KW   Chaperone; ATP-binding; Multigene family.
FT   CONFLICT     16     16       R -> T (IN REF. 1).
FT   CONFLICT     47     47       A -> D (IN REF. 1).
FT   CONFLICT     81     81       S -> T (IN REF. 1).
FT   CONFLICT    263    268       CPFEPP -> VHLNLL (IN REF. 1).
FT   CONFLICT    325    325       L -> I (IN REF. 1).
FT   CONFLICT    477    477       G -> N (IN REF. 1).
SQ   SEQUENCE   551 AA;  60676 MW;  B3B39B41ED05BF42 CRC64;


  Query Match             62.5%;  Score 35;  DB 1;  Length 551;
  Best Local Similarity   66.7%;  Pred. No. 47;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARVQVV 12
          ||:||||   ||
Db    299 IDDVKKAGADVV 310


RESULT  13
AX2E_PHAAU
ID   AX2E_PHAAU     STANDARD;      PRT;   203 AA.
AC   O24543;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Auxin-induced protein 22E (Indole-3-acetic acid induced protein
DE   ARG14).
```

```
GN   AUX22E OR ARG14.
OS   Phaseolus aureus (Mung bean) (Vigna radiata).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.
OX   NCBI_TaxID=3916;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Hypocotyl;
RA   Hashimoto H., Yamamoto K.T.;
RT   "Three more members of the Aux/IAA gene family from mung bean (Vigna
RT   radiata) hypocotyl.";
RL   (In) Plant Gene Register PGR97-137.
CC   -!- FUNCTION: COULD ACT AS REGULATOR OF GENES RESPONSIBLE FOR
CC       MEDIATING THE VARIOUS AUXIN-INDUCED EVENTS RESPONSIBLE FOR CELL
CC       GROWTH (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- INDUCTION: BY AUXIN.
CC   -!- SIMILARITY: BELONGS TO THE AUX/IAA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB004933; BAA20849.1; -.
DR   Mendel; 25713; Phaau;1524;25713.
DR   InterPro; IPR003311; AUX_IAA.
DR   Pfam; PF02309; AUX_IAA; 1.
SQ   SEQUENCE   203 AA;  22694 MW;  4B13504E5A191F49 CRC64;


  Query Match             60.7%;  Score 34;  DB 1;  Length 203;
  Best Local Similarity   63.6%;  Pred. No. 27;
  Matches    7;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 DNVKKARVQVV 12
          |:|: |:|||
Db     72 DSVQPAKVQVV 82


RESULT  14
VATE_SCHPO
ID   VATE_SCHPO     STANDARD;      PRT;   227 AA.
AC   O13687;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Vacuolar ATP synthase subunit E (EC 3.6.3.14) (V-ATPase E subunit)
DE   (Vacuolar proton pump E subunit) (V-ATPase 26 kDa subunit).
GN   SPAC11E3.07.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Murphy L., Harris D., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: SUBUNIT OF THE PERIPHERAL V1 COMPLEX OF VACUOLAR ATPASE
CC       ESSENTIAL FOR ASSEMBLY OR CATALYTIC FUNCTION. V-ATPASE IS
CC       RESPONSIBLE FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS
CC       IN EUKARYOTIC CELLS.
CC   -!- SUBUNIT: V-ATPase is an heteromultimeric enzyme composed of a
CC       peripheral catalytic V1 complex (components A to H) attached to
CC       an integral membrane V0 proton pore complex (components: a, c, c',
CC       c'' and d) (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE V-ATPASE E SUBUNIT FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z98595; CAB11186.1; -.
DR   InterPro; IPR002842; vATP-synt_E.
DR   Pfam; PF01991; vATP-synt_E; 1.
KW   Hydrolase; ATP synthesis; Hydrogen ion transport.
SQ   SEQUENCE   227 AA;  25741 MW;  B5180B4FC02080CB CRC64;


  Query Match             60.7%;  Score 34;  DB 1;  Length 227;
  Best Local Similarity   41.7%;  Pred. No. 30;
  Matches    5;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARVQVV 12
          :||  :||:::|
Db    193 VDNTLRARLEIV 204


RESULT  15
GRPE_SYNY3
ID   GRPE_SYNY3     STANDARD;      PRT;   249 AA.
AC   Q59978;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   GrpE protein (HSP-70 cofactor).
GN   GRPE OR SLL0057.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96127529; PubMed=8590279;
RA   Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA   Sugiura M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT   region from map positions 64% to 92% of the genome.";
RL   DNA Res. 2:153-166(1995).
CC   -!- FUNCTION: STIMULATES, JOINTLY WITH DNAJ, THE ATPASE ACTIVITY OF
CC       DNAK. HELPS TO RELEASE ADP FROM DNAK THUS ALLOWING DNAK TO RECYCLE
CC       MORE EFFICIENTLY (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE GRPE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D64001; BAA10291.1; -.
DR   HSSP; P09372; 1DKG.
DR   InterPro; IPR000740; GrpE.
DR   Pfam; PF01025; GrpE; 1.
DR   PRINTS; PR00773; GRPEPROTEIN.
DR   PROSITE; PS01071; GRPE; 1.
KW   Chaperone; Heat shock; Complete proteome.
SQ   SEQUENCE   249 AA;  27567 MW;  C69CB5D0EB39B248 CRC64;


  Query Match             60.7%;  Score 34;  DB 1;  Length 249;
  Best Local Similarity   45.5%;  Pred. No. 33;
  Matches    5;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy       1 IDNVKKARVQV 11
           :|| ::|| |:
Db     137 VDNFERARTQI 147
```

Search completed: July 1, 2002, 16:30:13
Job time: 622 sec

GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:29:45 ; Search time 75.26 Seconds
(without alignments)
27.584 Million cell updates/sec

Title: US-09-461-061A-3
Perfect score: 56
Sequence: 1 IDNVKKARVQVV 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 40 | 71.4 | 540 | 5 | O96220 | O96220 plasmodium |
| 2 | 39 | 69.6 | 309 | 17 | O58817 | O58817 pyrococcus |
| 3 | 38 | 67.9 | 234 | 2 | O32330 | O32330 clostridium |
| 4 | 38 | 67.9 | 423 | 11 | P70517 | P70517 rattus norv |
| 5 | 38 | 67.9 | 430 | 11 | Q63581 | Q63581 rattus norv |
| 6 | 38 | 67.9 | 432 | 11 | Q91XK5 | Q91xk5 mus musculu |
| 7 | 37 | 66.1 | 253 | 12 | Q91MS5 | Q91ms5 lumpy skin |
| 8 | 36 | 64.3 | 220 | 17 | Q971Y5 | Q971y5 sulfolobus |
| 9 | 36 | 64.3 | 544 | 5 | Q26858 | Q26858 trypanosoma |
| 10 | 36 | 64.3 | 585 | 2 | Q46508 | Q46508 desulfovibr |
| 11 | 35 | 62.5 | 151 | 10 | Q9FF22 | Q9ff22 arabidopsis |
| 12 | 35 | 62.5 | 222 | 16 | Q92D88 | Q92d88 listeria in |
| 13 | 35 | 62.5 | 264 | 16 | Q9X250 | Q9x250 thermotoga |
| 14 | 35 | 62.5 | 274 | 2 | Q93KZ5 | Q93kz5 burkholderi |
| 15 | 35 | 62.5 | 305 | 10 | Q9LGL3 | Q9lgl3 oryza sativ |
| 16 | 35 | 62.5 | 334 | 16 | Q930W1 | Q930w1 rhizobium m |
| 17 | 35 | 62.5 | 388 | 2 | Q9FD82 | Q9fd82 staphylococ |
| 18 | 35 | 62.5 | 388 | 2 | Q9FD76 | Q9fd76 staphylococ |
| 19 | 35 | 62.5 | 421 | 17 | Q97W60 | Q97w60 sulfolobus |
| 20 | 35 | 62.5 | 479 | 16 | Q9RX19 | Q9rx19 deinococcus |
| 21 | 35 | 62.5 | 488 | 5 | Q9TZ72 | Q9tz72 caenorhabdi |
| 22 | 35 | 62.5 | 552 | 16 | Q9A975 | Q9a975 caulobacter |
| 23 | 35 | 62.5 | 759 | 16 | Q92FH3 | Q92fh3 listeria in |
| 24 | 35 | 62.5 | 818 | 16 | Q99W78 | Q99w78 staphylococ |
| 25 | 35 | 62.5 | 881 | 10 | Q9M300 | Q9m300 arabidopsis |
| 26 | 35 | 62.5 | 1291 | 10 | Q9SFE1 | Q9sfe1 arabidopsis |
| 27 | 35 | 62.5 | 2342 | 5 | O01677 | O01677 bombyx mori |
| 28 | 34 | 60.7 | 158 | 16 | Q9ZEC3 | Q9zec3 rickettsia |
| 29 | 34 | 60.7 | 246 | 10 | Q39505 | Q39505 cylindrothe |
| 30 | 34 | 60.7 | 250 | 17 | Q9V208 | Q9v208 pyrococcus |
| 31 | 34 | 60.7 | 264 | 16 | Q9AB70 | Q9ab70 caulobacter |
| 32 | 34 | 60.7 | 277 | 2 | Q9ZA57 | Q9za57 staphylococ |
| 33 | 34 | 60.7 | 363 | 16 | O34833 | O34833 bacillus su |
| 34 | 34 | 60.7 | 380 | 17 | O59472 | O59472 pyrococcus |
| 35 | 34 | 60.7 | 385 | 5 | Q965Q5 | Q965q5 caenorhabdi |
| 36 | 34 | 60.7 | 398 | 17 | Q975X3 | Q975x3 sulfolobus |
| 37 | 34 | 60.7 | 622 | 10 | Q9CAW9 | Q9caw9 arabidopsis |
| 38 | 34 | 60.7 | 622 | 10 | Q945L6 | Q945l6 arabidopsis |
| 39 | 34 | 60.7 | 861 | 5 | Q9NE02 | Q9ne02 leishmania |
| 40 | 34 | 60.7 | 923 | 5 | Q9N389 | Q9n389 caenorhabdi |
| 41 | 34 | 60.7 | 1301 | 5 | Q9U982 | Q9u982 drosophila |
| 42 | 34 | 60.7 | 1301 | 5 | Q9V8H6 | Q9v8h6 drosophila |
| 43 | 33 | 58.9 | 115 | 16 | Q986E7 | Q986e7 rhizobium l |
| 44 | 33 | 58.9 | 126 | 10 | Q9C7P0 | Q9c7p0 arabidopsis |
| 45 | 33 | 58.9 | 150 | 16 | Q9X299 | Q9x299 thermotoga |

ALIGNMENTS

RESULT 1
O96220
ID O96220 PRELIMINARY; PRT; 540 AA.
AC O96220;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE T-COMPLEX PROTEIN 1 (HSP60 FOLD SUPERFAMILY).
GN PFB0635W.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99021743; PubMed=9804551;
RA Gardner M.J., Tettelin H., Carucci D.J., Cummings L.M., Aravind L.,
RA Koonin E.V., Shallom S., Mason T., Yu K., Fujii C., Pederson J.,
RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M.,
RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Science 282:1126-1132(1998).
DR EMBL; AE001407; AAC71916.1; -.
DR InterPro; IPR002423; TCP1_cpn60.
DR Pfam; PF00118; cpn60_TCP1; 1.
DR PRINTS; PR00304; TCOMPLEXTCP1.
KW ATP-binding; Chaperone.
SQ SEQUENCE 540 AA; 60790 MW; 05AF2CEB613CCEC7 CRC64;

Query Match 71.4%; Score 40; DB 5; Length 540;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 IDNVKKARVQVV 12
           ||| ||| | |:
Db     277 IDNFKKANVDVI 288
```

```
RESULT   2
O58817
ID   O58817      PRELIMINARY;      PRT;   309 AA.
AC   O58817;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HYPOTHETICAL 35.5 KDA PROTEIN PH1090.
GN   PH1090.
OS   Pyrococcus horikoshii.
OC   Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX   NCBI_TaxID=53953;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OT3;
RX   MEDLINE=98344137; PubMed=9679194;
RA   Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA   Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA   Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA   Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA   Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA   Masuchi Y., Shizuya H., Kikuchi H.;
RT   "Complete sequence and gene organization of the genome of a hyper-
RT   thermophilic archaebacterium, Pyrococcus horikoshii OT3.";
RL   DNA Res. 5:55-76(1998).
DR   EMBL; AP000004; BAA30189.1; -.
DR   InterPro; IPR002629; Methionine_synt.
DR   Pfam; PF01717; Methionine_synt; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   309 AA;  35509 MW;  88ABAA43D2C7522D CRC64;


  Query Match             69.6%;  Score 39;  DB 17;  Length 309;
  Best Local Similarity   54.5%;  Pred. No. 28;
  Matches    6;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 DNVKKARVQVV 12
          :|:||||::|:
Db     46 ENLKKARIKVI 56


RESULT   3
O32330
ID   O32330      PRELIMINARY;      PRT;   234 AA.
AC   O32330;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
GN   PLSD.
OS   Clostridium butyricum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1492;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98053837; PubMed=9393688;
RA   Heath R.J., Goldfine H., Rock C.O.;
RT   "A gene (plsD) from Clostridium butyricum that functionally
RT   substitutes for the sn-glycerol-3-phosphate acyltransferase gene
RT   (plsB) of Escherichia coli.";
RL   J. Bacteriol. 179:7257-7263(1997).
DR   EMBL; AF009362; AAC46006.1; -.
DR   InterPro; IPR002123; Acyltransferase.
DR   Pfam; PF01553; Acyltransferase; 1.
KW   Transferase; Acyltransferase.
SQ   SEQUENCE   234 AA;  26198 MW;  A4C15861A2995C37 CRC64;


  Query Match             67.9%;  Score 38;  DB 2;  Length 234;
  Best Local Similarity   72.7%;  Pred. No. 33;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARVQV 11
          |||||| |: |
Db     44 IDNVKKPRIFV 54


RESULT   4
P70517
ID   P70517      PRELIMINARY;      PRT;   423 AA.
AC   P70517;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MAJOR ACUTE PHASE ALPHA-1 PROTEIN PRECURSOR (FRAGMENT).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cole T.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85149311; PubMed=2579644;
RA   Cole T., Inglis A., Nagashima M., Schreiber G.;
RT   "Major acute-phase alpha(1)-protein in the rat: Structure, molecular
RT   cloning, and regulation of mRNA levels.";
RL   Biochem. Biophys. Res. Commun. 126:719-724(1985).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85127561; PubMed=2578992;
RA   Cole T., Inglis A.S., Roxburgh C.M., Howlett G.J., Schreiber G.;
RT   "Major acute phase alphal-protein of the rat is homologous to bovine
RT   kininogen and contains the sequence for bradykinin: its synthesis is
RT   regulated at the mRNA level.";
RL   FEBS Lett. 182:57-61(1985).
DR   EMBL; K02814; AAA41569.1; -.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   Pfam; PF00031; cystatin; 3.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     11       POTENTIAL.
FT   CHAIN        12    423       POTENTIAL.
FT   CHAIN       371    379       POTENTIAL.
SQ   SEQUENCE   423 AA;  46905 MW;  F9E8BD3198547949 CRC64;


  Query Match             67.9%;  Score 38;  DB 11;  Length 423;
  Best Local Similarity   75.0%;  Pred. No. 61;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARVQVV 12
          || ||||  |||
Db    293 IDTVKKATSQVV 304


RESULT   5
Q63581
ID   Q63581      PRELIMINARY;      PRT;   430 AA.
AC   Q63581;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RAT T-KININOGEN (T-KG).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90034172; PubMed=2806908;
RA   Anderson K.P., Croyle M.L., Lingrel J.B.;
RT   "Primary structure of a gene encoding rat T-kininogen.";
RL   Gene 81:119-128(1989).
DR   EMBL; M29090; AAA42251.1; -.
DR   EMBL; M29083; AAA42251.1; JOINED.
DR   EMBL; M29084; AAA42251.1; JOINED.
DR   EMBL; M29091; AAA42251.1; JOINED.
DR   EMBL; M29085; AAA42251.1; JOINED.
DR   EMBL; M29086; AAA42251.1; JOINED.
DR   EMBL; M29087; AAA42251.1; JOINED.
DR   EMBL; M29088; AAA42251.1; JOINED.
DR   EMBL; M29089; AAA42251.1; JOINED.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   Pfam; PF00031; cystatin; 3.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
SQ   SEQUENCE   430 AA;  47618 MW;  45508DEF4BDC978C CRC64;


  Query Match             67.9%;  Score 38;  DB 11;  Length 430;
  Best Local Similarity   75.0%;  Pred. No. 62;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARVQVV 12
          || ||||  |||
Db    300 IDTVKKATSQVV 311


RESULT   6
Q91XK5
ID   Q91XK5      PRELIMINARY;      PRT;   432 AA.
AC   Q91XK5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ADULT FEMALE PLACENTA CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE   CLONE:1600027I01, FULL INSERT SEQUENCE.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=PLACENTA;
RA   Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,
RA   Arakawa T., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA   Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA   Imotani K., Ishii Y., Itoh M., Izawa M., Kato H., Kawai J., Kojima Y.,
RA   Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Okazaki Y., Okido T.,
RA   Owa C., Saito H., Saito R., Sakai C., Sakai K., Sano H., Sasaki D.,
RA   Shibata K., Shibata Y., Shinagawa A., Shiraki T., Sogabe Y.,
RA   Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T., Tejima Y.,
RA   Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA   Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=PLACENTA;
RX   MEDLINE=21085660; PubMed=11217851;
RA   RIKEN FANTOM Consortium.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=PLACENTA;
RX   MEDLINE=99279253; PubMed=10349636;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=PLACENTA;
RX   MEDLINE=20499374; PubMed=11042159;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=PLACENTA;
RX   MEDLINE=20530913; PubMed=11076861;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
DR   EMBL; AK005547; BAB24115.1; -.
SQ   SEQUENCE   432 AA;  47898 MW;  91854EDA5284A16B CRC64;


  Query Match             67.9%;  Score 38;  DB 11;  Length 432;
  Best Local Similarity   75.0%;  Pred. No. 63;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARVQVV 12
          || ||||  |||
Db    300 IDTVKKATSQVV 311


RESULT   7
Q91MS5
ID   Q91MS5      PRELIMINARY;      PRT;   253 AA.
AC   Q91MS5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   LSDV087 MUTT MOTIF PUTATIVE GENE EXPRESSION REGULATOR.
GN   LSDV087.
OS   lumpy skin disease virus.
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Capripoxvirus.
OX   NCBI_TaxID=59509;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NEETHLING 2490;
RX   MEDLINE=21329495; PubMed=11435593;
RA   Tulman E.R., Afonso C.L., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "Genome of lumpy skin disease virus.";
RL   J. Virol. 75:7122-7130(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NEETHLING 2490;
RA   Tulman E.R., Afonso C.L., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF325528; AAK85048.1; -.
SQ   SEQUENCE   253 AA;  30004 MW;  22D2573CA0BEA3B2 CRC64;


  Query Match             66.1%;  Score 37;  DB 12;  Length 253;
  Best Local Similarity   50.0%;  Pred. No. 55;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARVQVV 12
          :|||||  : :|
```

```
Db     43 LDNVKKVSISIV 54


RESULT   8
Q971Y5
ID   Q971Y5      PRELIMINARY;      PRT;   220 AA.
AC   Q971Y5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL PROTEIN ST1243.
GN   ST1243.
OS   Sulfolobus tokodaii.
OC   Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX   NCBI_TaxID=111955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JCM 10545 / 7;
RX   PubMed=11572479;
RA   Kawarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA   Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA   Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA   Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA   Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA   Oshima T., Kikuchi H.;
RT   "Complete genome sequence of an aerobic thermoacidophilic
RT   Crenarchaeon, Sulfolobus tokodaii strain7.";
RL   DNA Res. 8:123-140(2001).
DR   EMBL; AP000985; BAB66284.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   220 AA;  24964 MW;  1776CA9DE0D54347 CRC64;


  Query Match             64.3%;  Score 36;  DB 17;  Length 220;
  Best Local Similarity   41.7%;  Pred. No. 72;
  Matches    5;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           :||::| |:: |
Db     175 LDNIRKNRIEAV 186


RESULT   9
Q26858
ID   Q26858      PRELIMINARY;      PRT;   544 AA.
AC   Q26858;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HEXOSE TRANSPORTER.
OS   Trypanosoma cruzi.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX   NCBI_TaxID=5693;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CL;
RX   MEDLINE=94336729; PubMed=8058795;
RA   Tetaud E., Bringaud F., Chabas S., Barret M., Baltz T.;
RT   "Characterization of glucose transport and cloning of a hexose
RT   transporter gene in Trypanosoma cruzi.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:8278-8282(1994).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR   EMBL; U05588; AAA21207.1; -.
DR   InterPro; IPR000504; RRM.
DR   InterPro; IPR003662; sub_transporter.
DR   InterPro; IPR003663; Sugar_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
KW   Transmembrane.
SQ   SEQUENCE   544 AA;  58733 MW;  395F69DC3DD0E8A9 CRC64;


  Query Match             64.3%;  Score 36;  DB 5;  Length 544;
  Best Local Similarity   58.3%;  Pred. No. 1.9e+02;
  Matches    7;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           ::|:| |:||||
Db      25 LENLKVAQVQVV 36


RESULT  10
Q46508
ID   Q46508      PRELIMINARY;      PRT;   585 AA.
AC   Q46508;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   POTENTIAL NAD-REDUCING HYDROGENASE SUBUNIT.
OS   Desulfovibrio fructosovorans.
OC   Bacteria; Proteobacteria; delta subdivision; Desulfovibrio.
OX   NCBI_TaxID=878;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95270577; PubMed=7751270;
RA   Malki S., Saimmaime I., De Luca G., Rousset M., Dermoun Z.,
RA   Belaich J.P.;
RT   "Characterization of an operon encoding an NADP-reducing hydrogenase
RT   in Desulfovibrio fructosovorans.";
RL   J. Bacteriol. 177:2628-2636(1995).
DR   EMBL; U07229; AAA87057.1; -.
DR   HSSP; P29166; 1FEH.
DR   InterPro; IPR001450; 4Fe4S_ferredoxin.
DR   InterPro; IPR001041; Ferredoxin.
DR   InterPro; IPR004108; Fe_hyd_lg_C.
DR   InterPro; IPR003149; Fe_hyd_SSU.
DR   Pfam; PF00111; fer2; 1.
DR   Pfam; PF00037; fer4; 1.
DR   Pfam; PF02906; Fe_hyd_lg_C; 1.
DR   Pfam; PF02256; Fe_hyd_SSU; 1.
DR   PRINTS; PR00353; 4FE4SFRDOXIN.
DR   PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.
KW   Iron-sulfur.
SQ   SEQUENCE   585 AA;  63429 MW;  C14D0EF6AA3E2A72 CRC64;


  Query Match             64.3%;  Score 36;  DB 2;  Length 585;
  Best Local Similarity   63.6%;  Pred. No. 2e+02;
  Matches    7;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQV 11
           :| ||||:|:|
Db     463 MDGVKKAKVKV 473


RESULT  11
Q9FF22
ID   Q9FF22      PRELIMINARY;      PRT;   151 AA.
AC   Q9FF22;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   GENOMIC DNA, CHROMOSOME 5, P1 CLONE:MXI10.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COLUMBIA;
RX   MEDLINE=97471969; PubMed=9330910;
RA   Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA   Miyajima N., Tabata S.;
```

```
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT   features of the 1.6 Mb regions covered by twenty physically assigned
RT   P1 clones.";
RL   DNA Res. 4:215-230(1997).
DR   EMBL; AB005248; BAB09352.1; -.
SQ   SEQUENCE   151 AA;  18044 MW;  C2C20F06F3385B3B CRC64;


  Query Match             62.5%;  Score 35;  DB 10;  Length 151;
  Best Local Similarity   50.0%;  Pred. No. 75;
  Matches    6;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARVQVV 12
          :|::|:|||  |
Db      3 VDSMKRARVSAV 14


RESULT  12
Q92D88
ID   Q92D88      PRELIMINARY;      PRT;   222 AA.
AC   Q92D88;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   LIN0929 PROTEIN.
GN   LIN0929.
OS   Listeria innocua.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Listeria.
OX   NCBI_TaxID=1642;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CLIP 11262 / SEROVAR 6A;
RX   PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
DR   EMBL; AL596166; CAC96161.1; -.
DR   ListiList; LIN00929; -.
KW   Complete proteome.
SQ   SEQUENCE   222 AA;  24626 MW;  CE9555C360C7CD47 CRC64;


  Query Match             62.5%;  Score 35;  DB 16;  Length 222;
  Best Local Similarity   50.0%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARVQVV 12
          ||| | ||: ::
Db    175 IDNTKDARITLI 186


RESULT  13
Q9X250
ID   Q9X250      PRELIMINARY;      PRT;   264 AA.
AC   Q9X250;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   HYPOTHETICAL 29.9 KDA PROTEIN.
GN   TM1727.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogales; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
DR   EMBL; AE001812; AAD36792.1; -.
DR   TIGR; TM1727; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   264 AA;  29883 MW;  3ABD9A420A4E7FE7 CRC64;


  Query Match             62.5%;  Score 35;  DB 16;  Length 264;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 DNVKKARVQ 10
          ||:|| ||:
Db    201 DNIKKMRVE 209


RESULT  14
Q93KZ5
ID   Q93KZ5      PRELIMINARY;      PRT;   274 AA.
AC   Q93KZ5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SCTJ.
GN   SCTJ.
OS   Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC   Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC   Burkholderia.
OX   NCBI_TaxID=28450;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=E503;
RX   MEDLINE=99330209; PubMed=10403415;
RA   Winstanley C., Hales B.A., Hart C.A.;
RT   "Evidence for the presence in Burkholderia pseudomallei of a type III
RT   secretion system-associated gene cluster.";
RL   J. Med. Microbiol. 48:649-656(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=E503;
RA   Winstanley C., Rainbow L., Hart C.A.;
RT   "Distribution of type III secretion gene clusters in Burkholderia
RT   pseudomallei, Burkholderia thailandensis and Burkholderia mallei.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF074878; AAK73236.1; -.
SQ   SEQUENCE   274 AA;  29653 MW;  727AE95D20EA5FA4 CRC64;


  Query Match             62.5%;  Score 35;  DB 2;  Length 274;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+02;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARVQVV 12
          || |  ||||:|
Db    130 IDGVLVARVQIV 141


RESULT  15
Q9LGL3
ID   Q9LGL3      PRELIMINARY;      PRT;   305 AA.
```

```
AC   Q9LGL3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SIMILAR TO ARABIDOPSIS THALIANA LECTIN RECEPTOR KINASE.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NIPPONBARE;
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT   clone:P0041E11.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NIPPONBARE;
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT   clone:P0433F09.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP002521; BAA96771.1; -.
DR   EMBL; AP002539; BAB08210.1; -.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00098; zf-CCHC; 1.
DR   SMART; SM00343; ZnF_C2HC; 1.
KW   Kinase; Lectin; Receptor; Zinc-finger.
SQ   SEQUENCE   305 AA;  33666 MW;  0E5CC16EB8EC2958 CRC64;


  Query Match             62.5%;  Score 35;  DB 10;  Length 305;
  Best Local Similarity   63.6%;  Pred. No. 1.6e+02;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 DNVKKARVQVV 12
           : ||||||| :
Db     112 ERVKKARVQTL 122


Search completed: July  1, 2002, 16:29:46
Job time: 700 sec
```

GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:19:43 ; Search time 95.97 Seconds
(without alignments)
13.889 Million cell updates/sec

Title: US-09-461-061A-3
Perfect score: 56
Sequence: 1 IDNVKKARVQVV 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 56 | 100.0 | 12 | 21 | AAY95407 | Anti-angiogenic pe |
| 2 | 56 | 100.0 | 16 | 21 | AAB08553 | Peptide identified |
| 3 | 56 | 100.0 | 26 | 18 | AAW54336 | Bradykinin analogo |
| 4 | 56 | 100.0 | 32 | 21 | AAY95408 | Anti-angiogenic D3 |
| 5 | 56 | 100.0 | 117 | 14 | AAR33350 | Domaine 3, bradyki |
| 6 | 56 | 100.0 | 122 | 21 | AAB37447 | Human kininogen D3 |
| 7 | 56 | 100.0 | 123 | 21 | AAY95426 | Human high mol.wt. |
| 8 | 56 | 100.0 | 644 | 22 | ABG21101 | Novel human diagno |
| 9 | 52 | 92.9 | 369 | 22 | ABG21099 | Novel human diagno |
| 10 | 52 | 92.9 | 435 | 22 | ABG21105 | Novel human diagno |
| 11 | 48 | 85.7 | 248 | 22 | ABG21102 | Novel human diagno |
| 12 | 44 | 78.6 | 434 | 5 | AAP40633 | Bradykinin protein |
| 13 | 44 | 78.6 | 436 | 5 | AAP40257 | Bradykinin protein |
| 14 | 40 | 71.4 | 160 | 22 | AAG82639 | S. epidermidis ope |
| 15 | 40 | 71.4 | 540 | 21 | AAB18269 | Plasmodium falcipa |
| 16 | 39 | 69.6 | 16 | 21 | AAY95410 | Anti-angiogenic D3 |
| 17 | 35 | 62.5 | 56 | 22 | AAM06553 | Human foetal prote |
| 18 | 35 | 62.5 | 370 | 22 | AAG81510 | S. epidermidis ope |
| 19 | 35 | 62.5 | 388 | 22 | AAB60851 | Mevalonate pathway |
| 20 | 35 | 62.5 | 388 | 22 | AAB60852 | Mevalonate pathway |
| 21 | 35 | 62.5 | 810 | 22 | AAU34090 | Staphylococcus aur |
| 22 | 35 | 62.5 | 817 | 22 | AAG81478 | S. epidermidis ope |
| 23 | 35 | 62.5 | 817 | 22 | AAG82217 | S. epidermidis ope |
| 24 | 35 | 62.5 | 818 | 22 | AAU36831 | Staphylococcus aur |
| 25 | 35 | 62.5 | 818 | 22 | AAU37196 | Staphylococcus aur |
| 26 | 34 | 60.7 | 262 | 22 | AAB88511 | Haemophilus influe |
| 27 | 34 | 60.7 | 583 | 22 | AAG82132 | S. epidermidis ope |
| 28 | 34 | 60.7 | 702 | 22 | AAU33479 | Enterococcus faeca |
| 29 | 34 | 60.7 | 704 | 22 | AAU35289 | Enterococcus faeca |
| 30 | 34 | 60.7 | 1268 | 22 | ABB61314 | Drosophila melanog |
| 31 | 34 | 60.7 | 1475 | 11 | AAR08221 | Recombinant alpha |
| 32 | 33 | 58.9 | 80 | 21 | AAB00168 | Brain specific mem |
| 33 | 33 | 58.9 | 208 | 22 | AAU38356 | Salmonella typhi c |
| 34 | 33 | 58.9 | 217 | 22 | AAU34655 | E. coli cellular p |
| 35 | 33 | 58.9 | 251 | 21 | AAB00167 | Brain specific mem |
| 36 | 33 | 58.9 | 274 | 21 | AAB14118 | Bordetella pertuss |
| 37 | 33 | 58.9 | 285 | 21 | AAG14565 | Arabidopsis thalia |
| 38 | 33 | 58.9 | 292 | 21 | AAG14564 | Arabidopsis thalia |
| 39 | 33 | 58.9 | 323 | 21 | AAG14563 | Arabidopsis thalia |
| 40 | 33 | 58.9 | 328 | 22 | AAG98392 | Escherichia coli p |
| 41 | 33 | 58.9 | 342 | 20 | AAY16785 | Human secreted pro |
| 42 | 33 | 58.9 | 342 | 21 | AAB00163 | Brain specific mem |
| 43 | 33 | 58.9 | 353 | 22 | AAM25782 | Human protein sequ |
| 44 | 33 | 58.9 | 392 | 21 | AAG28027 | Arabidopsis thalia |
| 45 | 33 | 58.9 | 409 | 21 | AAG28026 | Arabidopsis thalia |

ALIGNMENTS

RESULT 1
AAY95407
ID AAY95407 standard; Peptide; 12 AA.
XX
AC AAY95407;
XX
DT 25-SEP-2000 (first entry)
XX
DE Anti-angiogenic peptide C-terminal fragment.
XX
KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW therapy; human; D3 peptide.
XX
OS Homo sapiens.
XX
PN WO200035407-A2.
XX
PD 22-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28465.
XX
PR 16-DEC-1998; 98US-0112427.
XX
PA (UTEM ) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI McCrae RK;
XX
DR WPI; 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,

PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT 3 analog -
XX
PS Claim 3; Page 25; 44pp; English.
XX
CC The present sequence is that of a C-terminal fragment of a novel
CC anti-angiogenic D3 peptide (see AAY95408) derived from human high
CC mol.wt. kininogen (HK) domain 3 (see AAY95426). The full-length D3
CC peptide inhibits endothelial cell proliferation and thus possesses
CC anti-angiogenic activity. It is an example of peptides of the
CC invention (see AAY95405-26) that are analogues of certain sites in
CC the HK domain 3. The peptides inhibit endothelial cell proliferation
CC and may also induce endothelial cell apoptosis. Compositions
CC including the peptides are used in claimed methods for inhibiting
CC angiogenesis, inhibiting endothelial cell proliferation, and
CC inducing endothelial cell apoptosis. Cancer, rheumatoid arthritis,
CC and ocular disorders characterized by undesired vascularization of
CC the retina are treated.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 56; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IDNVKKARVQVV 12
||||||||||||
Db 1 idnvkkarvqvv 12

RESULT 2
AAB08553
ID AAB08553 standard; Peptide; 16 AA.
XX
AC AAB08553;
XX
DT 20-DEC-2000 (first entry)
XX
DE Peptide identified from an origin of prepro-bradykinine.
XX
KW Precursor peptide; polypeptide hormone; peptide identification.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "hydrogen attached"
FT Modified-site 16
FT /note= "amidated residue"
XX
PN WO200050636-A1.
XX
PD 31-AUG-2000.
XX
PF 24-FEB-2000; 2000WO-FR00460.
XX
PR 25-FEB-1999; 99US-0257525.
XX
PA (SCRC ) SCRAS SOC CONSEILS RECH & APPL SCI.
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Camara Ferrer YJA, Thurieau C, Martinez J, Berge G, Goze C;
XX
DR WPI; 2000-572101/53.
XX
PT Identifying peptide with selected function, useful particularly for
PT C-amidated hormones, by screening database for combination of nucleic
PT acid and amino acid sequences -
XX
PS Claim 16; Page 20; 40pp; French.
XX
CC The specification describes a method for identifying a peptide having
CC a particular function. The method comprises preparing a database of
CC polynucleotides and polypeptides of unknown functions, screening the
CC database for a combination of nucleotides or amino acids indicative of
CC the peptide with a particular function, and identifying polynucleotides
CC and proteins which contain the peptide. The method is used to identify
CC precursor peptides with an amidated C-terminus, especially polypeptide
CC hormones, for studying physiologically active substances. The present
CC sequence represents a peptide which was identified using the method of
CC the invention.
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 56; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IDNVKKARVQVV 12
||||||||||||
Db 4 idnvkkarvqvv 15

RESULT 3
AAW54336
ID AAW54336 standard; peptide; 26 AA.
XX
AC AAW54336;
XX
DT 30-JUL-1998 (first entry)
XX
DE Bradykinin analogous peptide 19.
XX
KW Inhibition; thrombin-induced platelet; prevention; platelet aggregation;
KW ADP-induced activation.
XX
OS Synthetic.
XX
PN WO9641640-A1.
XX
PD 27-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US09940.
XX
PR 09-JUN-1995; 95US-0000096.
XX
PA (UNMI ) UNIV MICHIGAN.
XX
PI Hasan AAK, Schmaier AH;
XX
DR WPI; 1997-065304/06.
XX
PT Inhibition of platelet activation and aggregation - by admin. of new
PT or known bradykinin analogues
XX
PS Disclosure; Page 44; 73pp; English.
XX
CC Administration of a peptide or multimer related to bradykinin or other
CC disclosed peptides and multimers can be used for the inhibition of
CC thrombin-induced platelets or other cells. They can also be used for
CC preventing platelet aggregation, or inhibiting ADP-induced activation.
CC This is useful to prevent arterial occlusions arising from coronary
CC thrombosis and stroke.
XX
SQ Sequence 26 AA;

Query Match 100.0%; Score 56; DB 18; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.00076;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IDNVKKARVQVV 12
||||||||||||
Db 8 idnvkkarvqvv 19

```
RESULT   4
AAY95408
ID   AAY95408 standard; Peptide; 32 AA.
XX
AC   AAY95408;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Anti-angiogenic D3 peptide.
XX
KW   Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW   endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW   rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW   therapy; human; D3 peptide.
XX
OS   Homo sapiens.
XX
PN   WO200035407-A2.
XX
PD   22-JUN-2000.
XX
PF   02-DEC-1999;   99WO-US28465.
XX
PR   16-DEC-1998;   98US-0112427.
XX
PA   (UTEM ) UNIV TEMPLE.
PA   (MCCR/) MCCRAE R K.
XX
PI   McCrae RK;
XX
DR   WPI; 2000-442247/38.
XX
PT   Composition for inhibiting angiogenesis and endothelial cell
PT   proliferation, inducing endothelial cell apoptosis and treating cancer,
PT   rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT   3 analog  -
XX
PS   Claim 4; Page 26; 44pp; English.
XX
CC   The present sequence is that of a D3 peptide derived from human
CC   high mol.wt. kininogen (HK) domain 3 (see AAY95426).  The D3 peptide
CC   inhibits endothelial cell proliferation and thus possesses
CC   anti-angiogenic activity.  It is an example of D3 peptides of the
CC   invention (see AAY95405-26) that are analogues of certain sites in
CC   the HK domain 3, in this case amino acid residues Asn275-Lys282.
CC   The peptides inhibit endothelial cell proliferation and may also
CC   induce endothelial cell apoptosis.  Compositions including the
CC   peptides are used in claimed methods for inhibiting angiogenesis,
CC   inhibiting endothelial cell proliferation, and inducing endothelial
CC   cell apoptosis.  Cancer, rheumatoid arthritis, and ocular disorders
CC   characterized by undesired vascularization of the retina are treated.
XX
SQ   Sequence   32 AA;


  Query Match             100.0%;  Score 56;  DB 21;  Length 32;
  Best Local Similarity   100.0%;  Pred. No. 0.00096;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           ||||||||||||
Db      21 idnvkkarvqvv 32


RESULT   5
AAR33350
ID   AAR33350 standard; protein; 117 AA.
XX
AC   AAR33350;
XX
DT   01-JUL-1993  (first entry)
XX
DE   Domaine 3, bradykinin release activating peptide.
XX
KW   Domain 3; human; kininogen; heavy chain; low molecular weight; plasma;
KW   trypsin; platelet; activation; granule contents; hemostasis; thrombin;
KW   tissue plasminogen activator;thrombosis; inflammatory response;
KW   endothelial cell; von Willebrand factor;
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..18
FT                   /note= "Leader peptide"
FT   Protein         19..117
FT                   /note= "Mature protein"
XX
PN   WO9303748-A.
XX
PD   04-MAR-1993.
XX
PF   13-AUG-1992;   92WO-US06809.
XX
PR   13-AUG-1991;   91US-0744545.
XX
PA   (UTEM ) UNIV TEMPLE.
XX
PI   Jiang Y,  Schmaier AB;
XX
DR   WPI; 1993-093714/11.
XX
PT   Use of trypsin-cleavage fragment of human kininogen - for
PT   increasing vascular bradykinin release, for lowering blood
PT   pressure and treating hypertension
XX
PS   Disclosure; Fig 1; 46pp; English.
XX
CC   The sequence given represents domain 3, amino acids 246-362, of
CC   the human kininogen heavy chain.  Domain 3 was isolated from low
CC   molecular weight kininogen, derived from human plasma, by cleavage
CC   with trypsin.  Domain 3 peptide inhibits platelet activation causing
CC   a marked decrease in the platelets ability to aggregate and secrete
CC   their granule contents.  The granule contents comprise proteins which
CC   participate in hemostasis, thrombosis and the inflammatory response.
CC   Domain 3 also inhibits endothelial cell activation shown by a decrease
CC   in secretion of endothelial cell contents such as tissue plasminogen
CC   activator and von Willebrand factor.  Domain 3 functions to inhibit
CC   cell activation by blocking thrombin binding to its target cells, the
CC   peptide is a selective inhibitor of thrombin-induced platlet
CC   activation.
XX
SQ   Sequence   117 AA;


  Query Match             100.0%;  Score 56;  DB 14;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 0.004;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           ||||||||||||
Db      38 idnvkkarvqvv 49


RESULT   6
AAB37447
ID   AAB37447 standard; protein; 122 AA.
XX
AC   AAB37447;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human kininogen D3.
XX
KW   Enzyme; legumain; endopeptidase; cystatin; human; kininogen.
```

XX
OS Homo sapiens.
XX
PN WO200064945-A1.
XX
PD 02-NOV-2000.
XX
PF 20-APR-2000; 2000WO-GB01571.
XX
PR 22-APR-1999; 99GB-0009133.
XX
PA (BABR-) BABRAHAM INST.
XX
PI Abrahamson M, Barrett AJ;
XX
DR WPI; 2000-687316/67.
XX
PT Inhibition of mammalian legumain or legumain-related endopeptidase by
PT cystatin involves interaction with second papain-non-reactive site of
PT cystatin -
XX
PS Disclosure; Fig 4; 45pp; English.
XX
CC The present invention relates to inhibition of the enzymatic activity of
CC legumain or a legumain-related endopeptidase by cystatin. The inhibition
CC involves an interaction between legumain and a papain-non-reactive site
CC of cystatin. Legumain (EC 3.4.22.34) is a cysteine endopeptidase, and
CC performs a protein-processing function. The present sequence is human
CC kininogen D3, which was used in the present invention. Kininogen is a
CC type 3 cystatin.
XX
SQ Sequence 122 AA;

Query Match 100.0%; Score 56; DB 21; Length 122;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IDNVKKARVQVV 12
||||||||||||
Db 43 idnvkkarvqvv 54

RESULT 7
AAY95426
ID AAY95426 standard; Peptide; 123 AA.
XX
AC AAY95426;
XX
DT 25-SEP-2000 (first entry)
XX
DE Human high mol.wt. kininogen domain 3.
XX
KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW therapy; human; D3 peptide.
XX
OS Homo sapiens.
XX
PN WO200035407-A2.
XX
PD 22-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28465.
XX
PR 16-DEC-1998; 98US-0112427.
XX
PA (UTEM ) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI McCrae RK;
XX
DR WPI; 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,
PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT 3 analog -
XX
PS Disclosure; Page 4; 44pp; English.
XX
CC The present sequence is that of domain 3 of human high mol.wt.
CC kininogen (HK). The invention provides peptides (see AAY95405-24)
CC that are analogues of certain sites in the HK domain 3,
CC specifically Asn275-Lys282, Cys246-Cys249, Leu331-Tyr338 and
CC Tyr299-Ser314. The peptides, in which native Cys residues may be
CC replaced by Ala residues, inhibit endothelial cell proliferation
CC and may also induce endothelial cell apoptosis. Compositions
CC including the peptides are used in claimed methods for inhibiting
CC angiogenesis, inhibiting endothelial cell proliferation, and
CC inducing endothelial cell apoptosis. Cancer, rheumatoid arthritis,
CC and ocular disorders characterized by undesired vascularization of
CC the retina are treated.
XX
SQ Sequence 123 AA;

Query Match 100.0%; Score 56; DB 21; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IDNVKKARVQVV 12
||||||||||||
Db 49 idnvkkarvqvv 60

RESULT 8
ABG21101
ID ABG21101 standard; Protein; 644 AA.
XX
AC ABG21101;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #21092.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS85288.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 51460; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 644 AA;

Query Match 100.0%; Score 56; DB 22; Length 644;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IDNVKKARVQVV 12
||||||||||||
Db 301 idnvkkarvqvv 312

RESULT 9
ABG21099
ID ABG21099 standard; Protein; 369 AA.
XX
AC ABG21099;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #21090.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS85286.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 51458; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 369 AA;

Query Match 92.9%; Score 52; DB 22; Length 369;
Best Local Similarity 100.0%; Pred. No. 0.081;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IDNVKKARVQV 11
|||||||||||
Db 341 idnvkkarvqv 351

RESULT 10
ABG21105
ID ABG21105 standard; Protein; 435 AA.
XX
AC ABG21105;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #21096.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS85292.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 51464; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

```
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   435 AA;


  Query Match           92.9%;  Score 52;  DB 22;  Length 435;
  Best Local Similarity  91.7%;  Pred. No. 0.097;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARVQVV 12
          |||||||||:||
Db    305 idnvkkarvrvv 316


RESULT  11
ABG21102
ID   ABG21102 standard; Protein; 248 AA.
XX
AC   ABG21102;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #21093.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS85289.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 51461; 103pp; English..
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   248 AA;


  Query Match           85.7%;  Score 48;  DB 22;  Length 248;
  Best Local Similarity  100.0%;  Pred. No. 0.3;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARVQ 10
          ||||||||||
Db     98 idnvkkarvq 107


RESULT  12
AAP40633
ID   AAP40633 standard; Protein; 434 AA.
XX
AC   AAP40633;
XX
DT   30-JUL-1992  (first entry)
XX
DE   Bradykinin protein precursor: type II (pKG146, pKG254).
XX
KW   Blood pressure; kininogen; probe..
XX
FH   Key             Location/Qualifiers
FT   Peptide         378..386
FT                   /label= bradykinin
FT   Peptide         391..395
FT                   /note= "probe (AAN40241)-encoded sequence"
XX
PN   JP59125896-A.
XX
PD   20-JUL-1984.
XX
PF   07-JAN-1983;   83JP-0000984.
XX
PR   07-JAN-1983;   83JP-0000984.
XX
PA   (MITU ) MITSUBISHI CHEM IND KK.
XX
DR   WPI; 1984-216122/35.
DR   N-PSDB; AAN40314.
XX
PT   c-Dna fragment of protein precursor - used to code bradykinin
XX
PS   Disclosure; Fig 2; 6 pp; Japanese.
XX
CC   Bradykinin is a peptide consisting of nine amino acids.  It has the
CC   biological effect of decreasing blood pressure.  Although kininogen
CC   is known as a protein-precursor of bradykinin, its structure is unknown
CC   because of the difficulty in collecting large enough samples of
CC   kininogen for structural investigation.
XX
SQ   Sequence   434 AA;
```

```
  Query Match             78.6%;  Score 44;  DB 5;  Length 434;
  Best Local Similarity   83.3%;  Pred. No. 3.2;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           || |||| ||||
Db     298 idtvkkatvqvv 309


RESULT  13
AAP40257
ID   AAP40257 standard; Protein; 436 AA.
XX
AC   AAP40257;
XX
DT   30-JUL-1992  (first entry)
XX
DE   Bradykinin protein precursor: type I (pKG13, pKG59).
XX
KW   Blood pressure; kininogen; probe.
XX
FH   Key             Location/Qualifiers
FT   Peptide         380..388
FT                   /label= bradykinin
FT   Peptide         393..397
FT                   /note= "probe (AAN40241)-encoded sequence"
XX
PN   JP59125896-A.
XX
PD   20-JUL-1984.
XX
PF   07-JAN-1983;   83JP-0000984.
XX
PR   07-JAN-1983;   83JP-0000984.
XX
PA   (MITU ) MITSUBISHI CHEM IND KK.
XX
DR   WPI; 1984-216122/35.
DR   N-PSDB; AAN40242.
XX
PT   c-Dna fragment of protein precursor - used to code bradykinin
XX
PS   Disclosure; Fig 2; 6 pp; Japanese.
XX
CC   Bradykinin is a peptide consisting of nine amino acids.  It has the
CC   biological effect of decreasing blood pressure.  Although kininogen
CC   is known as a protein-precursor of bradykinin, its structure is unknown
CC   because of the difficulty in collecting large enough samples of
CC   kininogen for structural investigation.
XX
SQ   Sequence   436 AA;


  Query Match             78.6%;  Score 44;  DB 5;  Length 436;
  Best Local Similarity   83.3%;  Pred. No. 3.2;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           || |||| ||||
Db     300 idtvkkatvqvv 311


RESULT  14
AAG82639
ID   AAG82639 standard; Protein; 160 AA.
XX
AC   AAG82639;
XX
DT   03-SEP-2001  (first entry)
XX
DE   S. epidermidis open reading frame protein sequence SEQ ID NO:2372.
XX
KW   Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW   vaccination; endocarditis.
XX
OS   Staphylococcus epidermidis.
XX
PN   WO200134809-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30782.
XX
PR   09-NOV-1999;   99US-0164258.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Kimmerly WJ;
XX
DR   WPI; 2001-316495/33.
DR   N-PSDB; AAH53489.
XX
PT   Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT   useful for vaccinating against infections, e.g. endocarditis -
XX
PS   Claim 18; Page 627; 2188pp; English.
XX
CC   AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC   (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC   (I) and (II) can have antibacterial activity and therefore can be used
CC   in vaccination. The nucleic acids (I) may be used to produce the
CC   S. epidermidis polypeptides (II) via the production of vectors
CC   containing them which are used to produce hosts cells which express the
CC   polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC   used to vaccinate subjects and to raise antibodies against the bacteria.
CC   The polypeptides may also be used to assay for other inhibitors of their
CC   activity and therefore identify compounds that may be used for the
CC   treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC   AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC   polynucleotide sequences from the present invention. AAH55091 to
CC   AAH55098 represent oligonucleotide sequences and primers which are used
CC   in the exemplification of the present invention.
CC   N.B. The present invention specifically claims all the polynucleotide
CC   sequences given in the sequence listing of the present specification,
CC   however the sequence listing only goes up to SEQ ID NO:4454 so even
CC   though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC   no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ   Sequence   160 AA;


  Query Match             71.4%;  Score 40;  DB 22;  Length 160;
  Best Local Similarity   75.0%;  Pred. No. 6.1;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           ||||||  | ||
Db     124 idnvkkttvyvv 135


RESULT  15
AAB18269
ID   AAB18269 standard; Protein; 540 AA.
XX
AC   AAB18269;
XX
DT   07-NOV-2000  (first entry)
XX
DE   Plasmodium falciparum chromosome 2 related protein SEQ ID NO:126.
XX
KW   Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
KW   antimalarial; malaria; protozoacide; infection; insecticide.
XX
OS   Plasmodium falciparum.
XX
```

PN WO200025728-A2.
XX
PD 11-MAY-2000.
XX.
PF 05-NOV-1999; 99WO-US26796.
XX
PR 05-NOV-1998; 98US-0107131.
XX
PA (HOFF/) HOFFMAN S.
PA (CARU/) CARUCCI D.
PA (GARD/) GARDNER M.
PA (VENT/) VENTER J C.
XX
PI Hoffman S, Carucci D, Gardner M, Venter JC;
XX
DR WPI; 2000-365347/31.
XX
PT Proteins encoded by chromosome 2 of the human malarial parasite,
PT Plasmodium falciparum, useful as antimalarial vaccines and in the
PT diagnosis of P.falciparum infection -
XX
PS Disclosure; Page 296-298; 577pp; English.
XX
CC The present invention describes proteins and their fragments (I) encoded
CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC Also described are: (1) nucleotide sequences (II) encoding (I); and (2)
CC vaccines against P. falciparum infection comprising (I) or (II).
CC (I) and (II) are useful for the development of vaccines against
CC P. falciparum infection. (I) and polyclonal antisera or a monoclonal
CC antibody raised to immunogens comprising the sequences of (I), are
CC useful in the detection of infection with P. falciparum. Furthermore,
CC (I) (especially when they are rifins or secreted or membrane proteins)
CC can aid the identification of drugs to treat or prevent P. falciparum
CC infection, or they can be used to identify drug resistance in
CC P. falciparum. Sequencing of the Plasmodium chromosome 2 and the
CC subsequent identification of proteins encoded by it will help to expand
CC our understanding of parasite biology, a process hampered by the
CC complexity of the parasitic lifecycle, and provide new targets for
CC vaccine and drug development. Parasite resistance to drugs and mosquito
CC resistance to insecticides have led to a resurgence of malaria in many
CC parts of the world, and there is a pressing need for vaccines and new
CC drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352 represent nucleotide
CC and protein sequences given in the present invention, but which are not
CC specifically mentioned within the specification.
XX
SQ Sequence 540 AA;

Query Match 71.4%; Score 40; DB 21; Length 540;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 IDNVKKARVQVV 12
             ||| ||| | |:
Db     277 idnfkkanvdvi 288
```

Search completed: July 1, 2002, 16:19:44
Job time: 148 sec

GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:17:16 ; Search time 46.58 Seconds
(without alignments)
16.503 Million cell updates/sec

Title: US-09-461-061A-1
Perfect score: 45
Sequence: 1 NNATFYFK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 45 | 100.0 | 427 | 1 | KGHUL1 | kininogen, LMW pre |
| 2 | 45 | 100.0 | 644 | 1 | KGHUH1 | kininogen, HMW pre |
| 3 | 39 | 86.7 | 433 | 2 | A28055 | K-kininogen, LMW I |
| 4 | 39 | 86.7 | 639 | 2 | A25486 | kininogen, HMW I p |
| 5 | 36 | 80.0 | 858 | 2 | T24062 | hypothetical prote |
| 6 | 36 | 80.0 | 1132 | 2 | T31107 | telomerase reverse |
| 7 | 35 | 77.8 | 182 | 2 | T24206 | hypothetical prote |
| 8 | 34 | 75.6 | 71 | 2 | B83803 | hypothetical prote |
| 9 | 34 | 75.6 | 182 | 2 | T06978 | ABA-induced plasma |
| 10 | 34 | 75.6 | 273 | 2 | F86924 | hypothetical prote |
| 11 | 34 | 75.6 | 319 | 2 | A86650 | rhamnosyltransfera |
| 12 | 34 | 75.6 | 440 | 2 | T11319 | NADH dehydrogenase |
| 13 | 34 | 75.6 | 497 | 2 | G96611 | probable cytochrom |
| 14 | 34 | 75.6 | 630 | 2 | T25830 | hypothetical prote |
| 15 | 34 | 75.6 | 758 | 2 | T31994 | hypothetical prote |
| 16 | 34 | 75.6 | 949 | 2 | H97322 | DNA/RNA helicase, |
| 17 | 33 | 73.3 | 303 | 2 | T32658 | hypothetical prote |
| 18 | 33 | 73.3 | 332 | 2 | T01483 | hypothetical prote |
| 19 | 33 | 73.3 | 469 | 2 | F69403 | hypothetical prote |
| 20 | 33 | 73.3 | 603 | 2 | S70849 | cholinesterase (EC |
| 21 | 33 | 73.3 | 706 | 2 | D84466 | hypothetical prote |
| 22 | 33 | 73.3 | 895 | 2 | T23191 | hypothetical prote |
| 23 | 32 | 71.1 | 484 | 2 | A40774 | phosphocholine-bin |
| 24 | 32 | 71.1 | 562 | 2 | AF0852 | secretory protein |
| 25 | 32 | 71.1 | 563 | 2 | S54420 | invasion protein i |
| 26 | 32 | 71.1 | 567 | 2 | E91095 | type III secretion |
| 27 | 32 | 71.1 | 567 | 2 | A85941 | type III secretion |
| 28 | 32 | 71.1 | 606 | 2 | T10982 | NADH dehydrogenase |
| 29 | 32 | 71.1 | 1215 | 2 | T25078 | hypothetical prote |
| 30 | 32 | 71.1 | 1659 | 1 | OKBYN2 | protein kinase GCN |
| 31 | 31 | 68.9 | 140 | 2 | T10059 | cytokinin-induced |
| 32 | 31 | 68.9 | 148 | 2 | B71529 | hypothetical prote |
| 33 | 31 | 68.9 | 171 | 2 | E90088 | hypothetical prote |
| 34 | 31 | 68.9 | 190 | 2 | F84366 | hypothetical prote |
| 35 | 31 | 68.9 | 196 | 2 | F84941 | amidotransferase h |
| 36 | 31 | 68.9 | 202 | 2 | AF0604 | probable permease |
| 37 | 31 | 68.9 | 211 | 2 | T34501 | hypothetical prote |
| 38 | 31 | 68.9 | 240 | 2 | A46179 | U2 snRNP auxiliary |
| 39 | 31 | 68.9 | 249 | 2 | F69409 | conserved hypothet |
| 40 | 31 | 68.9 | 262 | 2 | S59078 | conserved hypothet |
| 41 | 31 | 68.9 | 264 | 2 | JC6125 | U2 small nuclear r |
| 42 | 31 | 68.9 | 334 | 1 | RDECEP | N-acetyl-gamma-glu |
| 43 | 31 | 68.9 | 334 | 2 | G91239 | N-acetyl-gamma-glu |
| 44 | 31 | 68.9 | 334 | 2 | D86087 | N-acetyl-gamma-glu |
| 45 | 31 | 68.9 | 350 | 2 | A82299 | outer membrane pro |

ALIGNMENTS

RESULT 1
KGHUL1
kininogen, LMW precursor [validated] - human
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Homo sapiens (man)
C;Date: 06-Jul-1982 #sequence_revision 27-Nov-1985 #text_change 08-Dec-2000
C;Accession: A01280; B25276; A27900; A27699; A31905; A34030
R;Ohkubo, I.; Kurachi, K.; Takasawa, T.; Shiokawa, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A;Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its ide
A;Reference number: A90490; MUID:85122621
A;Accession: A01280
A;Molecule type: mRNA
A;Residues: 1-427 <OHK>
A;Cross-references: GB:K02566; NID:g177889; PIDN:AAA35497.1; PID:g177890
R;Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A;Title: Cloning and sequence analysis of cDNAs for human high molecular weight and l
A;Reference number: A92544; MUID:85234582
A;Accession: B25276
A;Molecule type: mRNA
A;Residues: 1-427 <TAK>
A;Cross-references: GB:M11437; NID:g186751; PIDN:AAB59551.1; PID:g386853
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Rauth, G.; Mueller-Esterl, W.
in Kinins IV, part A, Greenbaum, L.M., and Margolius, H.S., eds., pp.91-95, Plenum, N
A;Title: Amino acid sequence of the light chain of human low molecular mass kininogen
A;Reference number: A27900
A;Accession: A27900
A;Molecule type: protein
A;Residues: 390-427 <LOT>
R;Mindroiu, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
Biochem. Biophys. Res. Commun. 152, 519-526, 1988
A;Title: A new kinin moiety in human plasma kininogens.
A;Reference number: A27699; MUID:88209021
A;Accession: A27699
A;Molecule type: protein
A;Residues: 380-389 <MIN>
R;Maeda, H.; Matsumura, Y.; Kato, H.
J. Biol. Chem. 263, 16051-16054, 1988
A;Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic f
A;Reference number: A31905; MUID:89034061
A;Accession: A31905
A;Molecule type: protein
A;Residues: 381-389 <MAE>
R;Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
Biochem. Biophys. Res. Commun. 150, 511-516, 1988
A;Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human p
A;Reference number: A34030; MUID:88106632
A;Accession: A34030
A;Molecule type: protein
A;Residues: 380-389 <SAS>

R;Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 260, 8610-8617, 1985
A;Title: Structural organization of the human kininogen gene and a model for its evoluti
A;Reference number: A92545; MUID:85234583
A;Contents: annotation; gene organization
R;Pierce, J.V.
Fed. Proc. 27, 52-57, 1968
A;Title: Structural features of plasma kinins and kininogens.
A;Reference number: A91455; MUID:90255622
A;Contents: annotation; bradykinin
C;Comment: The LMW kininogen precursor is produced from the same gene as the HMW form (s
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Genetics:
A;Gene: GDB:KNG
A;Cross-references: GDB:125256; OMIM:228960
A;Map position: 3q27-3q27
A;Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3; 401/3
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; glyc
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-427/Product: LMW prokininogen (kininogen I) #status predicted <MAT>
F;19-389,390-427/Product: LMW kininogen II #status predicted <MAT2>
F;19-379/Product: LMW kininogen heavy chain #status predicted <HCH>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>
F;380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
F;390-427/Product: LMW kininogen light chain #status experimental <LCH>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;28-407,83-94,107-126,142-145,206-218,229-248,264-267,328-340,351-370/Disulfide bonds:
F;48,169,205,294/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
F;383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F;389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;401/Binding site: carbohydrate (Thr) (covalent) #status absent

Query Match 100.0%; Score 45; DB 1; Length 427;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 NNATFYFK 8
           ||||||||
Db     293 NNATFYFK 300
```

RESULT 2
KGHUH1
kininogen, HMW precursor [validated] - human
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen; prokininogen
N;Contains: bradykinin (kallidin I); HMW kininogen I; HMW kininogen II; low molecular we
C;Species: Homo sapiens (man)
C;Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 08-Dec-2000
C;Accession: A01279; A25276; S32422; A91153; A24871; A27899; A27699; A31905; A34030; S02
R;Ohkubo, I.; Kurachi, K.; Takasawa, T.; Shiokawa, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A;Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its identi
A;Reference number: A90490; MUID:85122621
A;Accession: A01279
A;Molecule type: mRNA
A;Residues: 1-389 <OHK>
A;Cross-references: GB:K02566; NID:g177889
R;Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A;Title: Cloning and sequence analysis of cDNAs for human high molecular weight and low
A;Reference number: A92544; MUID:85234582
A;Accession: A25276
A;Molecule type: mRNA
A;Residues: 1-592,'I',594-644 <TAK>
A;Cross-references: GB:M11437; NID:g186751; PIDN:AAB59550.1; PID:g386852
R;Auerswald, E.A.; Roessler, D.; Mentele, R.; Assfalg-Machleidt, I.
FEBS Lett. 321, 93-97, 1993
A;Title: Cloning, expression and characterization of human kininogen domain 3.
A;Reference number: S32422; MUID:93223854
A;Accession: S32422
A;Molecule type: mRNA
A;Residues: 'ANSM',253-377 <AUE>
A;Note: differences are due to known cloning artifacts
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A;Title: The amino acid sequence of the light chain of human high-molecular-mass kini
A;Reference number: A91153; MUID:86030270
A;Accession: A91153
A;Molecule type: protein
A;Residues: 379-644 <LOT>
A;Note: the bradykinin sequence preceding the light chain sequence was not determined
R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
Eur. J. Biochem. 154, 471-478, 1986
A;Title: Completion of the primary structure of human high-molecular-mass kininogen.
A;Reference number: A24871; MUID:86108361
A;Accession: A24871
A;Molecule type: protein
A;Residues: 'Z',20-380 <KEL1>
R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
in Kinins IV, Greenbaum, L.M., and Margolius, H.S., ed., pp.85-89, Plenum Press, New
A;Title: Amino acid sequence of the light chain of human high molecular mass kininoge
A;Reference number: A27899
A;Accession: A27899
A;Molecule type: protein
A;Residues: 379-389,'K',390-407,'Q',409-644 <KEL2>
R;Mindroiu, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
Biochem. Biophys. Res. Commun. 152, 519-526, 1988
A;Title: A new kinin moiety in human plasma kininogens.
A;Reference number: A27699; MUID:88209021
A;Accession: A27699
A;Molecule type: protein
A;Residues: 380-389 <MIN>
R;Maeda, H.; Matsumura, Y.; Kato, H.
J. Biol. Chem. 263, 16051-16054, 1988
A;Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic f
A;Reference number: A31905; MUID:89034061
A;Accession: A31905
A;Molecule type: protein
A;Residues: 381-389 <MAE>
R;Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
Biochem. Biophys. Res. Commun. 150, 511-516, 1988
A;Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human p
A;Reference number: A34030; MUID:88106632
A;Accession: A34030
A;Molecule type: protein
A;Residues: 380-389 <SAS>
R;Lenarcic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
Biol. Chem. Hoppe-Seyler 369, 257-261, 1988
A;Title: Human cathepsin B and cysteine proteinase inhibitors (CPIs) in inflammatory
A;Reference number: S02482; MUID:89076517
A;Accession: S02482
A;Molecule type: protein
A;Residues: 1-19;189-192;310-314;381-389 <LEN1>
R;Kato, H.; Matsumura, Y.; Maeda, H.
FEBS Lett. 232, 252-254, 1988
A;Title: Isolation and identification of hydroxyproline analogues of bradykinin in hu
A;Reference number: A61495; MUID:88211869
A;Accession: A61495
A;Molecule type: protein
A;Residues: 380-389 <KAT1>
A;Experimental source: urine
A;Note: this peptide had Pro-383 modified to 4-hydroxyproline
A;Accession: B61495
A;Molecule type: protein
A;Residues: 381-389 <KAT2>
A;Experimental source: urine
A;Note: this peptide had Pro-383 modified to 4-hydroxyproline
A;Accession: C61495

A;Molecule type: protein
A;Residues: 380-389 <KAT3>
R;Lenarcic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
FEBS Lett. 280, 211-215, 1991
A;Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
A;Reference number: S14303; MUID:91192133
A;Accession: S14447
A;Molecule type: protein
A;Residues: 264-359,'N',361-375 <LEN2>
R;Little, S.S.; Johnson, D.A.
Biochem. J. 307, 341-346, 1995
A;Title: Human mast cell tryptase isoforms: separation and examination of substrate-spec
A;Reference number: S55239; MUID:95251593
A;Accession: S55239
A;Molecule type: protein
A;Residues: 450-452,'X',454,'X',456 <LIT>
R;Straczek, J.; Maachi, F.; le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabet, P.; Bellevil
FEBS Lett. 373, 207-211, 1995
A;Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like
A;Reference number: S68059; MUID:96033974
A;Accession: S68059
A;Molecule type: protein
A;Residues: 431-434 <STR>
R;Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 260, 8610-8617, 1985
A;Title: Structural organization of the human kininogen gene and a model for its evoluti
A;Reference number: A92545; MUID:85234583
A;Contents: annotation; gene organization
R;Pierce, J.V.
Fed. Proc. 27, 52-57, 1968
A;Title: Structural features of plasma kinins and kininogens.
A;Reference number: A91455; MUID:90255622
A;Contents: annotation; bradykinin
C;Comment: The HMW kininogen precursor and the LMW form are produced from the same gene
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Genetics:
A;Gene: GDB:KNG
A;Cross-references: GDB:125256; OMIM:228960
A;Map position: 3q27-3q27
A;Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>
F;19-379,390-644/Product: HMW kininogen II #status experimental <MAT2>
F;19-379/Domain: HMW kininogen heavy chain #status experimental <HCH>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>
F;380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
F;390-644/Domain: HMW kininogen light chain #status experimental <LCH>
F;421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
F;431-434/Product: low molecular weight growth promoting factor #status experimental <GF
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;28-614,83-94,107-126,142-145,206-218,229-248,264-267,328-340,351-370/Disulfide bonds:
F;48/Binding site: carbohydrate (Asn) (covalent) #status absent
F;169,205,294/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
F;383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F;389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;401,533,542,546,557,571,593,628/Binding site: carbohydrate (Thr) (covalent) #status ex
F;577/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 45; DB 1; Length 644;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
||||||||
Db 293 NNATFYFK 300

RESULT 3
A28055
K-kininogen, LMW I precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 15-Nov-1996
C;Accession: A28055
R;Furuto-Kato, S.; Matsumoto, A.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 12054-12059, 1985
A;Title: Primary structures of the mRNAs encoding the rat precursors for bradykinin a
nhibitor.
A;Reference number: A92496; MUID:86008264
A;Accession: A28055
A;Molecule type: mRNA
A;Residues: 1-433 <FUR>
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-433/Product: K-kininogen, LMW I #status predicted <MAT>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>

Query Match 86.7%; Score 39; DB 2; Length 433;
Best Local Similarity 87.5%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
|| |||||
Db 293 NNHTFYFK 300

RESULT 4
A25486
kininogen, HMW I precursor - rat
N;Contains: bradykinin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Nov-1996
C;Accession: A25486
R;Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A;Title: Differing expression patterns and evolution of the rat kininogen gene family
A;Reference number: A92625; MUID:87137443
A;Accession: A25486
A;Molecule type: mRNA
A;Residues: 1-639 <KIT>
A;Note: the authors translated the codon CAA for residue 347 as Asn
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-639/Product: kininogen, HMW I #status predicted <MAT>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>

Query Match 86.7%; Score 39; DB 2; Length 639;
Best Local Similarity 87.5%; Pred. No. 7.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
|| |||||
Db 293 NNHTFYFK 300

RESULT 5
T24062
hypothetical protein R09A8.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans

```
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T24062
R;Wilkinson, J.
submitted to the EMBL Data Library, November 1995
A;Reference number: Z19836
A;Accession: T24062
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-858 <WIL>
A;Cross-references: EMBL:Z68009; PIDN:CAA92004.1; GSPDB:GN00028; CESP:R09A8.2
A;Experimental source: clone R09A8
C;Genetics:
A;Gene: CESP:R09A8.2
A;Map position: X
A;Introns: 197/1; 324/3; 377/3; 435/3; 495/3; 587/1; 627/3; 730/3; 812/1


  Query Match           80.0%;  Score 36;  DB 2;  Length 858;
  Best Local Similarity  75.0%;  Pred. No. 38;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 NNATFYFK 8
          | :|||||
Db     49 NESTFYFK 56


RESULT   6
T31107
telomerase reverse transciptase - Oxytricha trifallax
C;Species: Oxytricha trifallax
C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C;Accession: T31107
R;Bryan, T.M.; Sperger, J.M.; Chapman, K.B.; Cech, T.R.
Proc. Natl. Acad. Sci. U.S.A. 95, 8479-8484, 1998
A;Title: Telomerase reverse transcriptase genes in Tetrahymena thermophila and Oxytricha
A;Reference number: Z20985; MUID:98337940
A;Accession: T31107
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1132 <BRY>
A;Cross-references: EMBL:AF060230; NID:g3342795; PID:g3342796; PIDN:AAC39163.1
C;Genetics:
A;Gene: TERT


  Query Match           80.0%;  Score 36;  DB 2;  Length 1132;
  Best Local Similarity  75.0%;  Pred. No. 50;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 NNATFYFK 8
          || :||||
Db   1001 NNISFYFK 1008


RESULT   7
T24206
hypothetical protein R12H7.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 08-Dec-2000
C;Accession: T24206
R;Coles, L.
submitted to the EMBL Data Library, August 1995
A;Reference number: Z19854
A;Accession: T24206
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-182 <WIL>
A;Cross-references: EMBL:Z50755; PIDN:CAA90635.1; GSPDB:GN00028; CESP:R12H7.3
A;Experimental source: clone R12H7
C;Genetics:
A;Gene: CESP:R12H7.3
A;Map position: X
A;Introns: 150/3
C;Superfamily: human S-phase kinase-associated protein 1A


  Query Match           77.8%;  Score 35;  DB 2;  Length 182;
  Best Local Similarity  75.0%;  Pred. No. 13;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 NNATFYFK 8
          |||| :||
Db    161 NNATLFFK 168


RESULT   8
B83803
hypothetical protein BH1226 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: B83803
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: B83803
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-71 <STO>
A;Cross-references: GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BAB04945.1; GSPDB:G
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1226


  Query Match           75.6%;  Score 34;  DB 2;  Length 71;
  Best Local Similarity  100.0%;  Pred. No. 8.1;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 NATFYF 7
          ||||||
Db     43 NATFYF 48


RESULT   9
T06978
ABA-induced plasma membrane protein PM 19 - wheat
C;Species: Triticum aestivum (common wheat)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 08-Oct-1999
C;Accession: T06978
R;Koike, M.; Takezawa, D.; Arakawa, K.; Yoshida, S.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z15842
A;Accession: T06978
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-182 <KOI>
A;Cross-references: EMBL:U80037; NID:g1724111; PIDN:AAB38504.1; PID:g1724112
A;Experimental source: cv. Chihoku
C;Genetics:
A;Note: WTABAPM


  Query Match           75.6%;  Score 34;  DB 2;  Length 182;
  Best Local Similarity  85.7%;  Pred. No. 20;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 NNATFYF 7
          | |||||
Db     47 NGATFYF 53


RESULT  10
F86924
hypothetical protein [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
```

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: F86924
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: F86924
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-273 <STO>
A;Cross-references: GB:AL450380; NID:g13092504; PIDN:CAC29634.1; GSPDB:GN00147
C;Genetics:
A;Gene: ML0126

Query Match 75.6%; Score 34; DB 2; Length 273;
Best Local Similarity 85.7%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NNATFYF 7
|||.|||
Db 20 NNAQFYF 26

RESULT 11
A86650
rhamnosyltransferase [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: A86650
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrli
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: A86650
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-319 <STO>
A;Cross-references: GB:AE005176; PID:g12723056; PIDN:AAK04299.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: rgpB

Query Match 75.6%; Score 34; DB 2; Length 319;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NNATFYF 7
|||.|||
Db 80 NNADFYF 86

RESULT 12
T11319
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 2 - Pedinomonas minor mitochondrion
C;Species: mitochondrion Pedinomonas minor
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 17-Mar-2000
C;Accession: T11319
R;Turmel, M.; Lemieux, C.; Burger, G.; Lang, B.F.; Otis, C.; Plante, I.; Gray, M.W.
submitted to the EMBL Data Library, December 1998
A;Description: The complete mitochondrial DNA sequences of Nephroselmis olivacea and Ped
A;Reference number: Z17261
A;Accession: T11319
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-440 <TUR>
A;Cross-references: EMBL:AF116775; NID:g4378766; PID:g4378773; PIDN:AAD19671.1
C;Genetics:
A;Gene: nad2
A;Genome: mitochondrion
A;Genetic code: SGC3
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 2
C;Keywords: mitochondrion; NAD; oxidoreductase

Query Match 75.6%; Score 34; DB 2; Length 440;
Best Local Similarity 85.7%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NNATFYF 7
|||.|||
Db 238 NNALFYF 244

RESULT 13
G96611
probable cytochrome P450 T8L23.21 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: G96611
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719
A;Accession: G96611
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-497 <STO>
A;Cross-references: GB:AE005173; NID:g11055848; PIDN:AAG28316.1; GSPDB:GN00141
C;Genetics:
A;Gene: T8L23.21
A;Map position: 1
C;Superfamily: Candida cytochrome P450 52A3; cytochrome P450 homology

Query Match 75.6%; Score 34; DB 2; Length 497;
Best Local Similarity 85.7%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NATFYFK 8
.|.|||||
Db 64 NLTFYFK 70

RESULT 14
T25830
hypothetical protein M01A10.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T25830
R;Scheet, P.
submitted to the EMBL Data Library, February 1997
A;Description: The sequence of C. elegans cosmid M01A10.
A;Reference number: Z20094
A;Accession: T25830
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-630 <SCH>
A;Cross-references: EMBL:U88174; PIDN:AAB42276.1; GSPDB:GN00019; CESP:M01A10.4
A;Experimental source: strain Bristol N2; clone M01A10
C;Genetics:
A;Gene: CESP:M01A10.4
A;Map position: 1
A;Introns: 10/2; 249/3; 284/1

Query Match 75.6%; Score 34; DB 2; Length 630;
Best Local Similarity 75.0%; Pred. No. 68;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 NNATFYFK 8
           :|| |||| 
Db     444 SNAQFYFK 451
```

RESULT 15
T31994
hypothetical protein C49D10.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T31994
R;Henkhaus, J.; Wohldmann, P.; Beck, C.
submitted to the EMBL Data Library, July 1997
A;Description: The sequence of C. elegans cosmid C49D10.
A;Reference number: Z21108
A;Accession: T31994
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-758 <HEN>
A;Cross-references: EMBL:AF016665; PIDN:AAC71186.1; GSPDB:GN00020; CESP:C49D10.1
A;Experimental source: strain Bristol N2; clone C49D10
C;Genetics:
A;Gene: CESP:C49D10.1
A;Map position: 2
A;Introns: 438/2

Query Match 75.6%; Score 34; DB 2; Length 758;
Best Local Similarity 75.0%; Pred. No. 82;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 NNATFYFK 8
           :|| |||| 
Db     598 SNAQFYFK 605
```

Search completed: July 1, 2002, 16:20:36
Job time: 200 sec

GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:19:51 ; Search time 21.51 Seconds
(without alignments)
14.401 Million cell updates/sec

Title: US-09-461-061A-1
Perfect score: 45
Sequence: 1 NNATFYFK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 45 | 100.0 | 644 | 1 | KNG_HUMAN | P01042 homo sapien |
| 2 | 39 | 86.7 | 639 | 1 | KNG_RAT | P08934 rattus norv |
| 3 | 36 | 80.0 | 1132 | 1 | TERT_OXYTR | O76332 oxytricha t |
| 4 | 33 | 73.3 | 603 | 1 | CHLE_MOUSE | Q03311 mus musculu |
| 5 | 32 | 71.1 | 562 | 1 | INVG_SALTY | P35672 salmonella |
| 6 | 32 | 71.1 | 1590 | 1 | GCN2_YEAST | P15442 saccharomyc |
| 7 | 31 | 68.9 | 82 | 1 | U2AG_PIG | Q29350 sus scrofa |
| 8 | 31 | 68.9 | 196 | 1 | HIS5_BUCAI | P57204 buchnera ap |
| 9 | 31 | 68.9 | 239 | 1 | U2AG_MOUSE | Q9d883 mus musculu |
| 10 | 31 | 68.9 | 240 | 1 | U2AG_HUMAN | Q01081 homo sapien |
| 11 | 31 | 68.9 | 264 | 1 | U2AG_DROME | Q94535 drosophila |
| 12 | 31 | 68.9 | 334 | 1 | ARGC_ECOLI | P11446 escherichia |
| 13 | 31 | 68.9 | 341 | 1 | OMPU_VIBCH | P97085 vibrio chol |
| 14 | 31 | 68.9 | 434 | 1 | KNL2_BOVIN | P01047 bos taurus |
| 15 | 31 | 68.9 | 441 | 1 | PAFA_HUMAN | Q13093 h platelet- |
| 16 | 31 | 68.9 | 444 | 1 | PAFA_CANFA | Q28262 c platelet- |
| 17 | 31 | 68.9 | 619 | 1 | KNH2_BOVIN | P01045 bos taurus |
| 18 | 31 | 68.9 | 623 | 1 | RSD1_YEAST | P32368 saccharomyc |
| 19 | 31 | 68.9 | 901 | 1 | VEF_GVPU | P41723 pseudalatia |
| 20 | 31 | 68.9 | 901 | 1 | VEF_GVTN | P29998 trichoplusi |
| 21 | 31 | 68.9 | 902 | 1 | VEF_GVHA | P54232 heliothis a |
| 22 | 31 | 68.9 | 988 | 1 | ST23_YEAST | Q06010 saccharomyc |
| 23 | 31 | 68.9 | 1066 | 1 | NUC2_NEUCR | Q01317 neurospora |
| 24 | 31 | 68.9 | 2710 | 1 | TOXA_CLODI | P16154 clostridium |
| 25 | 31 | 68.9 | 4725 | 1 | DYHC_DICDI | P34036 dictyosteli |
| 26 | 30 | 66.7 | 163 | 1 | SFAS_ECOLI | P13430 escherichia |
| 27 | 30 | 66.7 | 196 | 1 | HIS5_METJA | Q57929 methanococc |
| 28 | 30 | 66.7 | 208 | 1 | LEUD_BUCDN | O85073 buchnera ap |
| 29 | 30 | 66.7 | 237 | 1 | YS76_CAEEL | Q09619 caenorhabdi |
| 30 | 30 | 66.7 | 259 | 1 | KKA6_ACIBA | P09885 acinetobact |
| 31 | 30 | 66.7 | 331 | 1 | NIXA_HELPJ | Q9zm74 helicobacte |
| 32 | 30 | 66.7 | 331 | 1 | NIXA_HELPY | Q48262 helicobacte |
| 33 | 30 | 66.7 | 365 | 1 | NMPC_ECOLI | P21420 escherichia |
| 34 | 30 | 66.7 | 383 | 1 | CYB_APILI | P34845 apis mellif |
| 35 | 30 | 66.7 | 397 | 1 | YAK7_SCHPO | Q09919 schizosacch |
| 36 | 30 | 66.7 | 452 | 1 | SMP1_YEAST | P38128 saccharomyc |
| 37 | 30 | 66.7 | 477 | 1 | GLGA_STRPN | Q97qs5 streptococc |
| 38 | 30 | 66.7 | 492 | 1 | CPBJ_MOUSE | O55071 mus musculu |
| 39 | 30 | 66.7 | 537 | 1 | YIV9_YEAST | P40583 saccharomyc |
| 40 | 30 | 66.7 | 541 | 1 | UL21_VZVD | P09289 varicella-z |
| 41 | 30 | 66.7 | 581 | 1 | CHLE_RABIT | P21927 oryctolagus |
| 42 | 30 | 66.7 | 602 | 1 | CHLE_HUMAN | P06276 homo sapien |
| 43 | 30 | 66.7 | 603 | 1 | NU5M_MYOGL | O63908 myoxus glis |
| 44 | 30 | 66.7 | 758 | 1 | L2DT_DROME | Q24371 drosophila |
| 45 | 30 | 66.7 | 952 | 1 | PM16_CHLPN | Q9z882 chlamydia p |

ALIGNMENTS

RESULT 1
KNG_HUMAN
ID KNG_HUMAN STANDARD; PRT; 644 AA.
AC P01042; P01043;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Kininogen precursor (Alpha-2-thiol proteinase inhibitor) [Contains:
DE Bradykinin].
GN KNG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC TISSUE=Liver;
RX MEDLINE=85234582; PubMed=2989293;
RA Takagaki Y., Kitamura N., Nakanishi S.;
RT "Cloning and sequence analysis of cDNAs for human high molecular
RT weight and low molecular weight prekininogens. Primary structures of
RT two human prekininogens.";
RL J. Biol. Chem. 260:8601-8609(1985).
RN [2]
RP GENE STRUCTURE.
RX MEDLINE=85234583; PubMed=2989294;
RA Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T.,
RA Nakanishi S.;
RT "Structural organization of the human kininogen gene and a model for
RT its evolution.";
RL J. Biol. Chem. 260:8610-8617(1985).
RN [3]
RP SEQUENCE OF 1-401 FROM N.A.
RX MEDLINE=85122621; PubMed=6441591;
RA Ohkubo I., Kurachi K., Takasawa T., Shiokawa H., Sasaki M.;
RT "Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and
RT its identity with low molecular weight kininogen.";
RL Biochemistry 23:5691-5697(1984).
RN [4]
RP SEQUENCE OF 379-644.
RX MEDLINE=86030270; PubMed=4054110;
RA Lottspeich F., Kellermann J., Henschen A., Foertsch B.,
RA Mueller-Esterl W.;
RT "The amino acid sequence of the light chain of human high-molecular-
RT mass kininogen.";
RL Eur. J. Biochem. 152:307-314(1985).
RN [5]
RP SEQUENCE OF 381-389.
RX MEDLINE=90255622; PubMed=4952632;
RA Pierce J.V.;
RT "Structural features of plasma kinins and kininogens.";
RL Fed. Proc. 27:52-57(1968).
RN [6]
RP DISULFIDE BONDS.
RA Sueyoshi T., Miyata T., Kato H., Iwanaga S.;
RT "Disulfide bonds in bovine HMW kininogens.";

```
RL   Seikagaku 56:808-808(1984).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC       HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
CC       FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-
CC       INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
CC       BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
CC       PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
CC       CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
CC       DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
CC       MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
CC       PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
CC       OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
CC       A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
CC       INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
CC       KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
CC       CLOTTING.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; K02566; AAA35497.1; -.
DR   EMBL; M11437; AAB59550.1; -.
DR   EMBL; M11438; AAB59550.1; JOINED.
DR   EMBL; M11521; AAB59550.1; JOINED.
DR   EMBL; M11522; AAB59550.1; JOINED.
DR   EMBL; M11523; AAB59550.1; JOINED.
DR   EMBL; M11524; AAB59550.1; JOINED.
DR   EMBL; M11525; AAB59550.1; JOINED.
DR   EMBL; M11526; AAB59550.1; JOINED.
DR   EMBL; M11527; AAB59550.1; JOINED.
DR   EMBL; M11528; AAB59550.1; JOINED.
DR   EMBL; M11437; AAB59551.1; -.
DR   EMBL; M11438; AAB59551.1; JOINED.
DR   EMBL; M11521; AAB59551.1; JOINED.
DR   EMBL; M11522; AAB59551.1; JOINED.
DR   EMBL; M11523; AAB59551.1; JOINED.
DR   EMBL; M11524; AAB59551.1; JOINED.
DR   EMBL; M11525; AAB59551.1; JOINED.
DR   EMBL; M11526; AAB59551.1; JOINED.
DR   EMBL; M11527; AAB59551.1; JOINED.
DR   EMBL; M11528; AAB59551.1; JOINED.
DR   PIR; A01279; KGHUH1.
DR   PIR; A25276; A25276.
DR   PIR; A01280; KGHUL1.
DR   PIR; B25276; B25276.
DR   PIR; S02482; S02482.
DR   SWISS-2DPAGE; P01042; HUMAN.
DR   MIM; 228960; -.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW   Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW   Alternative splicing.
FT   SIGNAL       1     18
FT   CHAIN        19    644       KININOGEN.
FT   CHAIN        19    380       KININOGEN HEAVY CHAIN.
FT   PEPTIDE      381   389       BRADYKININ.
FT   CHAIN        390   644       KININOGEN LIGHT CHAIN.
FT   DOMAIN       19    136       CYSTATIN-LIKE 1.
FT   DOMAIN       137   258       CYSTATIN-LIKE 2.
FT   DOMAIN       259   380       CYSTATIN-LIKE 3.
FT   DOMAIN       420   510       HIS-RICH (ASSOCIATED WITH CLOTTING
FT                                ACTIVITY).
FT   REPEAT       420   449
FT   REPEAT       450   479
FT   REPEAT       480   510
FT   MOD_RES      19    19        PYRROLIDONE CARBOXYLIC ACID.
FT   DISULFID     28    614       INTERCHAIN.
FT   DISULFID     83    94
FT   DISULFID     107   126
FT   DISULFID     142   145
FT   DISULFID     206   218
FT   DISULFID     229   248
FT   DISULFID     264   267
FT   DISULFID     328   340
FT   DISULFID     351   370
FT   CARBOHYD     48    48        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     169   169       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     205   205       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     294   294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     401   401       O-LINKED.
FT   CARBOHYD     533   533       O-LINKED.
FT   CARBOHYD     542   542       O-LINKED.
FT   CARBOHYD     546   546       O-LINKED.
FT   CARBOHYD     557   557       O-LINKED.
FT   CARBOHYD     571   571       O-LINKED.
FT   CARBOHYD     577   577       O-LINKED.
FT   CARBOHYD     593   593       O-LINKED.
FT   CARBOHYD     628   628       O-LINKED.
FT   VARSPLIC     402   427       VSPPHTSMAPAQDEERDSGKEQGHTR -> SHLRSCEYKGR
FT                                PPKAGAEPASEREVS (IN ISOFORM LMW).
FT   VARSPLIC     428   644       MISSING (IN ISOFORM LMW).
FT   CONFLICT     593   593       T -> I (IN REF. 1).
SQ   SEQUENCE   644 AA;  71945 MW;  3132B4CBAF8FBB7E CRC64;


  Query Match             100.0%;  Score 45;  DB 1;  Length 644;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NNATFYFK 8
           ||||||||
Db     293 NNATFYFK 300


RESULT   2
KNG_RAT
ID   KNG_RAT        STANDARD;      PRT;   639 AA.
AC   P08934; P08933;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Kininogen precursor [Contains: Bradykinin].
GN   KNG.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RX   MEDLINE=87137443; PubMed=3029068;
RA   Kitagawa H., Kitamura N., Hayashida H., Miyata T., Nakanishi S.;
RT   "Differing expression patterns and evolution of the rat kininogen
RT   gene family.";
RL   J. Biol. Chem. 262:2190-2198(1987).
RN   [2]
RP   SEQUENCE FROM N.A. (LMW ISOFORM).
```

```
RX   MEDLINE=86008264; PubMed=2413018;
RA   Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT   "Primary structures of the mRNAs encoding the rat precursors for
RT   bradykinin and T-kinin. Structural relationship of kininogens with
RT   major acute phase protein and alpha 1-cysteine proteinase
RT   inhibitor.";
RL   J. Biol. Chem. 260:12054-12059(1985).
RN   [3]
RP   SEQUENCE OF 1-65 FROM N.A.
RC   STRAIN=BUFFALO;
RX   MEDLINE=87250580; PubMed=2439509;
RA   Fung W.-P., Schreiber G.;
RT   "Structure and expression of the genes for major acute phase alpha 1-
RT   protein (thiostatin) and kininogen in the rat.";
RL   J. Biol. Chem. 262:9298-9308(1987).
RN   [4]
RP   SEQUENCE OF 1-41 FROM N.A.
RC   STRAIN=WISTAR; TISSUE=Liver;
RX   MEDLINE=87137465; PubMed=3818598;
RA   Kageyama R., Kitamura N., Ohkubo H., Nakanishi S.;
RT   "Differing utilization of homologous transcription initiation sites
RT   of rat K and T kininogen genes under inflammation condition.";
RL   J. Biol. Chem. 262:2345-2351(1987).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC       HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
CC       FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-
CC       INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
CC       BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
CC       PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
CC       CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
CC       DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
CC       MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
CC       PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
CC       OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
CC       A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
CC       INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
CC       KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
CC       CLOTTING.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- MISCELLANEOUS: RAT EXPRESS FOUR TYPES OF KININOGENS: THE CLASSICAL
CC       HMW/LMW KININOGENS AND TWO ADDITIONAL LMW-LIKE KININOGENS: T-I AND
CC       T-II.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29428; AAA41486.1; -.
DR   EMBL; M11884; AAA41487.1; -.
DR   EMBL; M14369; AAA41484.1; -.
DR   EMBL; M14369; AAA41485.1; ALT_SEQ.
DR   EMBL; M16455; AAA41482.1; -.
DR   PIR; A25486; A25486.
DR   PIR; A28055; A28055.
DR   HSSP; P01040; 1DVD.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW   Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW   Alternative splicing; Multigene family.
FT   SIGNAL       1    18
FT   CHAIN       19   639       KININOGEN.
FT   CHAIN       19   380       KININOGEN HEAVY CHAIN.
FT   PEPTIDE    381   389       BRADYKININ.
FT   CHAIN      390   639       KININOGEN LIGHT CHAIN.
FT   DOMAIN      19   136       CYSTATIN-LIKE 1.
FT   DOMAIN     137   258       CYSTATIN-LIKE 2.
FT   DOMAIN     259   380       CYSTATIN-LIKE 3.
FT   DOMAIN     439   514       HIS-RICH.
FT   DISULFID    28   609       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    83    94       BY SIMILARITY.
FT   DISULFID   107   126       BY SIMILARITY.
FT   DISULFID   142   145       BY SIMILARITY.
FT   DISULFID   206   218       BY SIMILARITY.
FT   DISULFID   229   248       BY SIMILARITY.
FT   DISULFID   264   267       BY SIMILARITY.
FT   DISULFID   328   340       BY SIMILARITY.
FT   DISULFID   351   370       BY SIMILARITY.
FT   CARBOHYD    82    82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   127   127       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   169   169       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   205   205       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   294   294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   529   529       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   402   433       VSPSYIARVQEERDPGNEQGPIHGHGWLHAKQ -> RLLNS
FT                              CEYKGRLLKAGAGPAPERQAEASTVTP (IN ISOFORM
FT                              LMW).
FT   VARSPLIC   434   639       MISSING (IN ISOFORM LMW).
FT   CONFLICT    61    61       E -> K (IN REF. 2).
SQ   SEQUENCE   639 AA;  70933 MW;  D3172DF94FF56AF5 CRC64;


  Query Match             86.7%;  Score 39;  DB 1;  Length 639;
  Best Local Similarity   87.5%;  Pred. No. 3.3;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 NNATFYFK 8
           || ||||| 
Db     293 NNHTFYFK 300


RESULT   3
TERT_OXYTR
ID   TERT_OXYTR     STANDARD;      PRT;  1132 AA.
AC   O76332;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Telomerase reverse transcriptase (EC 2.7.7.-) (Telomerase catalytic
DE   subunit) (Telomerase subunit P133).
GN   TERT.
OS   Oxytricha trifallax.
OC   Eukaryota; Alveolata; Ciliophora; hypotrichs; Stichotrichida;
OC   Oxytrichidae; Oxytricha.
OX   NCBI_TaxID=5946;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98337940; PubMed=9671703;
RA   Bryan T.M., Sperger J.M., Chapman K.B., Cech T.R.;
RT   "Telomerase reverse transcriptase genes identified in Tetrahymena
RT   thermophila and Oxytricha trifallax.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:8479-8484(1998).
CC   -!- FUNCTION: TELOMERASE IS A RIBONUCLEOPROTEIN ENZYME ESSENTIAL FOR
CC       THE REPLICATION OF CHROMOSOME TERMINI IN MOST EUKARYOTES. IT
CC       ELONGATES TELOMERES. IT IS A REVERSE TRANSCRIPTASE THAT ADDS
CC       SIMPLE SEQUENCE REPEATS TO CHROMOSOME ENDS BY COPYING A TEMPLATE
CC       SEQUENCE WITHIN THE RNA COMPONENT OF THE ENZYME.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: BELONGS TO THE REVERSE TRANSCRIPTASE FAMILY.
```

```
CC       TELOMERASE SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF060230; AAC39163.1; -.
DR   InterPro; IPR003545; Telomerase_RT.
DR   PRINTS; PR01365; TELOMERASERT.
KW   Transferase; RNA-directed DNA polymerase; Telomere; Nuclear protein;
KW   DNA-binding.
SQ   SEQUENCE   1132 AA;  134124 MW;  81E145F5F24392DC CRC64;


  Query Match             80.0%;  Score 36;  DB 1;  Length 1132;
  Best Local Similarity   75.0%;  Pred. No. 22;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 NNATFYFK 8
          || :||||
Db   1001 NNISFYFK 1008


RESULT   4
CHLE_MOUSE
ID   CHLE_MOUSE     STANDARD;      PRT;   603 AA.
AC   Q03311;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Cholinesterase precursor (EC 3.1.1.8) (Acylcholine acylhydrolase)
DE   (Choline esterase II) (Butyrylcholine esterase)
DE   (Pseudocholinesterase).
GN   BCHE.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90380429; PubMed=2400605;
RA   Rachinsky T.L., Camp S., Li Y., Ekstroem T.J., Newton M., Taylor P.;
RT   "Molecular cloning of mouse acetylcholinesterase: tissue distribution
RT   of alternatively spliced mRNA species.";
RL   Neuron 5:317-327(1990).
RN   [2]
RP   SEQUENCE OF 97-237 FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=91201348; PubMed=2016308;
RA   Arpagaus M., Chatonnet A., Masson P., Newton M., Vaughan T.A.,
RA   Bartels C.F., Nogueira C.P., la Du B.N., Lockridge O.;
RT   "Use of the polymerase chain reaction for homology probing of
RT   butyrylcholinesterase from several vertebrates.";
RL   J. Biol. Chem. 266:6966-6974(1991).
CC   -!- CATALYTIC ACTIVITY: An acylcholine + H(2)O = choline + a
CC       carboxylic acid anion.
CC   -!- SUBUNIT: HOMOTETRAMER. THE TETRAMER IS COMPOSED
CC       OF TWO DIMERS. THE TWO SUBUNITS IN A DIMER ARE LINKED BY A
CC       DISULFIDE BOND.
CC   -!- TISSUE SPECIFICITY: PRESENT IN MOST CELLS (EXCEPT ERYTHROCYTES).
CC   -!- MISCELLANEOUS: CHOLINESTERASE IS HIGHLY REACTIVE WITH
CC       ORGANOPHOSPHATE ESTERS.
CC   -!- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M99492; AAA37328.1; -.
DR   PIR; A39768; A39768.
DR   HSSP; P21836; 1MAH.
DR   MGD; MGI:894278; Bche.
DR   InterPro; IPR002018; Carboxylesterase_B.
DR   InterPro; IPR000997; Cholinesterase.
DR   InterPro; IPR000379; Est_lip_thioest_actsite.
DR   Pfam; PF00135; COesterase; 1.
DR   PRINTS; PR00878; CHOLNESTRASE.
DR   PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR   PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
KW   Hydrolase; Serine esterase; Glycoprotein; Signal.
FT   SIGNAL       1     29
FT   CHAIN       30    603       CHOLINESTERASE.
FT   ACT_SITE    227    227       BY SIMILARITY.
FT   ACT_SITE    354    354       BY SIMILARITY.
FT   ACT_SITE    467    467       BY SIMILARITY.
FT   DISULFID     94    121       BY SIMILARITY.
FT   DISULFID    281    292       BY SIMILARITY.
FT   DISULFID    429    548       BY SIMILARITY.
FT   DISULFID    600    600       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD     86     86       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    135    135       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    270    270       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    370    370       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    484    484       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    510    510       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    515    515       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    129    129       R -> P (IN REF. 2).
SQ   SEQUENCE   603 AA;  68521 MW;  719B1B220D1E5367 CRC64;


  Query Match             73.3%;  Score 33;  DB 1;  Length 603;
  Best Local Similarity   85.7%;  Pred. No. 44;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 NNATFYF 7
          ||| |||
Db    443 NNAFFYF 449


RESULT   5
INVG_SALTY
ID   INVG_SALTY     STANDARD;      PRT;   562 AA.
AC   P35672;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   InvG protein precursor.
GN   INVG OR STM2898.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TML;
RX   MEDLINE=95172719; PubMed=7868245;
RA   Lodge J., Douce G.R., Amin I.I., Bolton A.J., Martin G.D.,
RA   Chatfield S., Dougan G., Brown N.L., Stephen J.;
RT   "Biological and genetic characterization of TnphoA mutants of
RT   Salmonella typhimurium TML in the context of gastroenteritis.";
RL   Infect. Immun. 63:762-769(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SR-11;
RX   MEDLINE=95089692; PubMed=7997169;
RA   Kaniga K., Bossio J.C., Galan J.E.;
RT   "The Salmonella typhimurium invasion genes invF and invG encode
RT   homologues of the AraC and PulD family of proteins.";
```

```
RL   Mol. Microbiol. 13:555-568(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
CC   -!- FUNCTION: INVOLVED IN THE INVASION OF THE CELLS OF THE INTESTINAL
CC       EPITHELIUM. COULD BE NECESSARY FOR THE EXPORT OF INVASION RELATED
CC       DETERMINANTS.
CC   -!- SUBCELLULAR LOCATION: Outer membrane (Potential).
CC   -!- SIMILARITY: BELONGS TO THE PULD/OUTD/EXED/XPSD FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X75302; CAA53049.1; -.
DR   EMBL; U08280; AAA74040.1; -.
DR   EMBL; AE008832; AAL21778.1; -.
DR   PIR; S38422; S38422.
DR   StyGene; SG10188; invG.
DR   InterPro; IPR000016; Bac_GSPprotein.
DR   InterPro; IPR003522; SecIII_OMPG.
DR   Pfam; PF00263; GSPII_III; 1.
DR   PRINTS; PR01337; TYPE3OMGPROT.
DR   PROSITE; PS00875; T2SP_D; 1.
KW   Virulence; Transport; Protein transport; Signal; Outer membrane;
KW   Complete proteome.
FT   SIGNAL        1     14       POTENTIAL.
FT   CHAIN        15    562       INVG PROTEIN.
FT   CONFLICT     12     12       A -> R (IN REF. 2).
FT   CONFLICT    121    121       E -> Q (IN REF. 2).
FT   CONFLICT    197    205       LRDQKMVIP -> CAIRKWLFR (IN REF. 2).
FT   CONFLICT    232    240       AMPAFSANG -> RCQRFQRM (IN REF. 2).
FT   CONFLICT    243    243       G -> S (IN REF. 2).
FT   CONFLICT    262    264       AAA -> KPAEQ (IN REF. 2).
FT   CONFLICT    328    328       S -> T (IN REF. 1 AND 2).
FT   CONFLICT    329    329       I -> V (IN REF. 2).
FT   CONFLICT    370    380       RPVLLTQENVP -> APGITSSGKCS (IN REF. 2).
SQ   SEQUENCE   562 AA;  61765 MW;  8022905BE256058D CRC64;


  Query Match             71.1%;  Score 32;  DB 1;  Length 562;
  Best Local Similarity   75.0%;  Pred. No. 63;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 NNATFYFK 8
          || ||| |
Db    385 NNRTFYTK 392


RESULT   6
GCN2_YEAST
ID   GCN2_YEAST     STANDARD;      PRT;  1590 AA.
AC   P15442;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Protein kinase GCN2 (EC 2.7.1.-).
GN   GCN2 OR AAS1 OR YDR283C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89282814; PubMed=2660141;
RA   Wek R.C., Jackson B.M., Hinnebusch A.G.;
RT   "Juxtaposition of domains homologous to protein kinases and histidyl-
RT   tRNA synthetases in GCN2 protein suggests a mechanism for coupling
RT   GCN4 expression to amino acid availability.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:4579-4583(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE=88261291; PubMed=3290651;
RA   Roussou I., Thireos G., Hauge B.M.;
RT   "Transcriptional-translational regulatory circuit in Saccharomyces
RT   cerevisiae which involves the GCN4 transcriptional activator and the
RT   GCN2 protein kinase.";
RL   Mol. Cell. Biol. 8:2132-2139(1988).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RA   Le T.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   ASSOCIATION WITH RIBOSOMES.
RX   MEDLINE=91246169; PubMed=2038314;
RA   Ramirez M., Wek R.C., Hinnebusch A.G.;
RT   "Ribosome association of GCN2 protein kinase, a translational
RT   activator of the GCN4 gene of Saccharomyces cerevisiae.";
RL   Mol. Cell. Biol. 11:3027-3036(1991).
RN   [5]
RP   ACTIVITY ON SUI2.
RX   MEDLINE=92154672; PubMed=1739968;
RA   Dever T.E., Feng L., Wek R.C., Cigan A.M., Donahue T.D.,
RA   Hinnebusch A.G.;
RT   "Phosphorylation of initiation factor 2 alpha by protein kinase GCN2
RT   mediates gene-specific translational control of GCN4 in yeast.";
RL   Cell 68:585-596(1992).
CC   -!- FUNCTION: STIMULATES GCN4 TRANSLATION IN AMINO ACID-STARVED CELLS
CC       BY PHOSPHORYLATING THE ALPHA SUBUNIT OF EIF-2 (SUI2) ON SER-52.
CC   -!- SUBCELLULAR LOCATION: ASSOCIATES WITH THE 60S SUBUNIT OF
CC       RIBOSOMES.
CC   -!- INDUCTION: BY AMINO ACID STARVATION AND IN THE PRESENCE OF THE
CC       GCN4 PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       GCN2 SUBFAMILY.
CC   -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN IN POSITIONS 301
CC       TO 364 AND FROM POSITION 981 ONWARD DUE TO FRAMESHIFTS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M27082; AAA34636.1; -.
DR   EMBL; M20487; AAA34881.1; ALT_FRAME.
DR   EMBL; U51030; AAB64461.1; ALT_INIT.
DR   PIR; S05781; OKBYN2.
DR   PIR; A27723; A27723.
DR   HSSP; P24941; 1JST.
DR   SGD; S0002691; GCN2.
DR   InterPro; IPR002106; AA_tRNA_ligase_II.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; pkinase; 4.
DR   PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 2.
```

```
KW   Transferase; Serine/threonine-protein kinase; ATP-binding;
KW   Amino-acid biosynthesis; Aminoacyl-tRNA synthetase; Repeat.
FT   DOMAIN      187    458       PROTEIN KINASE 1.
FT   DOMAIN      530    912       PROTEIN KINASE 2.
FT   DOMAIN      930   1450       HISTIDYL-TRNA SYNTHETASE LIKE.
FT   NP_BIND     536    544       ATP (BY SIMILARITY).
FT   BINDING     559    559       ATP (BY SIMILARITY).
FT   ACT_SITE    766    766       BY SIMILARITY.
FT   CONFLICT    151    157       TIKAKLP -> NYKGKIA (IN REF. 2).
FT   CONFLICT    202    210       LMSSEMMEN -> YVFSNHGKS (IN REF. 2).
FT   CONFLICT    406    406       I -> M (IN REF. 2).
FT   CONFLICT    520    520       P -> A (IN REF. 2).
FT   CONFLICT    523    536       RSRYASDFEEIAVL -> DHDMLLTLKRLHFS
FT                                (IN REF. 2).
FT   CONFLICT    553    553       S -> T (IN REF. 2).
FT   CONFLICT    571    572       IL -> MI (IN REF. 2).
FT   CONFLICT    658    658       F -> C (IN REF. 2).
FT   CONFLICT    770    770       M -> K (IN REF. 2).
FT   CONFLICT    885    885       E -> Q (IN REF. 2).
SQ   SEQUENCE   1590 AA;  182041 MW;  768815424F64F1AA CRC64;


  Query Match             71.1%;  Score 32;  DB 1;  Length 1590;
  Best Local Similarity   75.0%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 NNATFYFK 8
          ||||| :|
Db    266 NNATFVWK 273


RESULT   7
U2AG_PIG
ID   U2AG_PIG       STANDARD;      PRT;    82 AA.
AC   Q29350;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Splicing factor U2AF 35 kDa subunit (U2 auxiliary factor 35 kDa
DE   subunit) (U2 snRNP auxiliary factor small subunit) (Fragment).
GN   U2AF1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Small intestine;
RX   MEDLINE=96327607; PubMed=8672129;
RA   Winteroe A.K., Fredholm M., Davies W.;
RT   "Evaluation and characterization of a porcine small intestine cDNA
RT   library: analysis of 839 clones.";
RL   Mamm. Genome 7:509-517(1996).
CC   -!- FUNCTION: PLAYS A CRITICAL ROLE IN BOTH CONSTITUTIVE AND ENHANCER-
CC       DEPENDENT SPLICING BY MEDIATING ESSENTIAL PROTEIN-PROTEIN
CC       INTERACTIONS AND PROTEIN-RNA INTERACTIONS REQUIRED FOR ACCURATE 3'
CC       SPLICE SITE SELECTION. DIRECTLY MEDIATES INTERACTIONS BETWEEN
CC       U2AF65 (LARGE SUBUNIT) AND PROTEINS BOUND TO THE ENHANCERS AND
CC       THUS MAY FUNCTION AS A BRIDGE BETWEEN U2AF65 AND THE ENHANCER
CC       COMPLEX TO RECRUIT IT TO THE ADJACENT INTRON (BY SIMILARITY).
CC   -!- SUBUNIT: ASSOCIATES WITH A 65 kDa PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC   -!- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; F14757; CAA23231.1; -.
DR   InterPro; IPR000571; Zf-CCCH.
DR   Pfam; PF00642; zf-CCCH; 1.
KW   Nuclear protein; RNA-binding; mRNA splicing; Zinc-finger; Repeat.
FT   DOMAIN       47     >82       INVOLVED IN BINDING U2AF-65 (BY
FT                                 SIMILARITY).
FT   NON_TER      82     82
SQ   SEQUENCE   82 AA;  9416 MW;  96882458044823A8 CRC64;


  Query Match             68.9%;  Score 31;  DB 1;  Length 82;
  Best Local Similarity   71.4%;  Pred. No. 14;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 NATFYFK 8
            | :||||
Db     17 NCSFYFK 23


RESULT   8
HIS5_BUCAI
ID   HIS5_BUCAI     STANDARD;      PRT;   196 AA.
AC   P57204;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Amidotransferase hisH (EC 2.4.2.-).
GN   HISH OR BU103.
OS   Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS   symbiotic bacterium).
OC   Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX   NCBI_TaxID=118099;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOKYO 1998;
RX   MEDLINE=20445173; PubMed=10993077;
RA   Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT   "Genome sequence of the endocellular bacterial symbiont of aphids
RT   Buchnera sp. APS.";
RL   Nature 407:81-86(2000).
CC   -!- FUNCTION: CATALYZES AN AMIDOTRANSFERASE REACTION THAT GENERATES
CC       IMIDAZOLE-GLYCEROL PHOSPHATE AND 5-AMINOIMIDAZOL-4-CARBOXAMIDE
CC       RIBONUCLEOTIDE, WHICH IS USED FOR PURINE SYNTHESIS.
CC   -!- PATHWAY: FIFTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE HISH FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP001118; BAB12822.1; -.
DR   InterPro; IPR000991; GATase_1.
DR   Pfam; PF00117; GATase; 1.
DR   PROSITE; PS00442; GATASE_TYPE_I; FALSE_NEG.
KW   Histidine biosynthesis; Transferase; Glutamine amidotransferase;
KW   Complete proteome.
FT   ACT_SITE     77     77       BY SIMILARITY.
FT   ACT_SITE    178    178       BY SIMILARITY.
FT   ACT_SITE    180    180       BY SIMILARITY.
SQ   SEQUENCE   196 AA;  21679 MW;  904624B74BBEBE5D CRC64;


  Query Match             68.9%;  Score 31;  DB 1;  Length 196;
  Best Local Similarity   71.4%;  Pred. No. 34;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 NNATFYF 7
```

```
          ||: |||
Db    135 NNSRFYF 141


RESULT   9
U2AG_MOUSE
ID   U2AG_MOUSE     STANDARD;      PRT;   239 AA.
AC   Q9D883; Q9CZ98; Q99LX2;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Splicing factor U2AF 35 kDa subunit (U2 auxiliary factor 35 kDa
DE   subunit) (U2 snRNP auxiliary factor small subunit).
GN   U2AF1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Small intestine;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [2]
RP   SEQUENCE OF 2-239 FROM N.A.
RC   TISSUE=Breast tumor;
RA   Strausberg R.;
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PLAYS A CRITICAL ROLE IN BOTH CONSTITUTIVE AND ENHANCER-
CC       DEPENDENT SPLICING BY MEDIATING ESSENTIAL PROTEIN-PROTEIN
CC       INTERACTIONS AND PROTEIN-RNA INTERACTIONS REQUIRED FOR ACCURATE 3'
CC       SPLICE SITE SELECTION. DIRECTLY MEDIATES INTERACTIONS BETWEEN
CC       U2AF65 (LARGE SUBUNIT) AND PROTEINS BOUND TO THE ENHANCERS AND
CC       THUS MAY FUNCTION AS A BRIDGE BETWEEN U2AF65 AND THE ENHANCER
CC       COMPLEX TO RECRUIT IT TO THE ADJACENT INTRON (By similarity).
CC   -!- SUBUNIT: ASSOCIATES WITH A 65 kDa PROTEIN (By similarity).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (By similarity).
CC   -!- DOMAIN: THE AMINO-TERMINAL SR-RICH DOMAIN IS REQUIRED FOR
CC       INTERACTIONS WITH SR PROTEINS AND THE SPLICING REGULATORS TRA AND
CC       TRA2, AND THE CARBOXY-TERMINAL DOMAIN IS REQUIRED FOR FORMATION OF
CC       THE U2AF35/U2AF65 HETERODIMER (By similarity).
CC   -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC   -!- SIMILARITY: CONTAINS 2 C3H1-TYPE ZINC FINGERS.
CC   -!- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AK008332; BAB25609.1; -.
DR   EMBL; AK012849; BAB28511.1; -.
DR   EMBL; BC002184; AAH02184.1; -.
DR   MGD; MGI:98884; U2af1.
DR   InterPro; IPR000504; RRM.
DR   InterPro; IPR003954; RRM_1.
DR   InterPro; IPR000571; Zf-CCCH.
DR   Pfam; PF00076; rrm; 1.
DR   Pfam; PF00642; zf-CCCH; 2.
DR   SMART; SM00360; RRM; 1.
DR   SMART; SM00361; RRM_1; 1.
DR   SMART; SM00356; ZnF_C3H1; 2.
DR   PROSITE; PS50102; RRM; 1.
DR   PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW   Nuclear protein; RNA-binding; mRNA splicing; Zinc-finger; Repeat.
FT   ZN_FING      13     41       C3H1-TYPE 1.
FT   DOMAIN       65    147       RNA-BINDING (RRM).
FT   ZN_FING     149    173       C3H1-TYPE 2.
FT   DOMAIN      179    238       ARG/GLY/SER-RICH (RS DOMAIN).
FT   CONFLICT    187    187       G -> R (IN REF. 1; BAB25609).
SQ   SEQUENCE   239 AA;  27815 MW;  DFF944210581244D CRC64;


  Query Match             68.9%;  Score 31;  DB 1;  Length 239;
  Best Local Similarity   71.4%;  Pred. No. 41;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 NATFYFK 8
           | :|||
Db      17 NCSFYFK 23


RESULT  10
U2AG_HUMAN
ID   U2AG_HUMAN     STANDARD;      PRT;   240 AA.
AC   Q01081;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Splicing factor U2AF 35 kDa subunit (U2 auxiliary factor 35 kDa
DE   subunit) (U2 snRNP auxiliary factor small subunit).
GN   U2AF1 OR U2AF35.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 68-89 AND 126-151.
RC   TISSUE=Fetal brain;
RX   MEDLINE=92409598; PubMed=1388271;
RA   Zhang M., Zamore P.D., Carmo-Fonseca M., Lamond A.I., Green M.R.;
RT   "Cloning and intracellular localization of the U2 small nuclear
RT   ribonucleoprotein auxiliary factor small subunit.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:8769-8773(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20289799; PubMed=10830953;
RA   Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA   Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA   Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA   Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA   Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA   Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA   Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA   Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA   Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA   Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA   Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA   Lehrach H., Reinhardt R., Yaspo M.-L.;
RT   "The DNA sequence of human chromosome 21.";
RL   Nature 405:311-319(2000).
RN   [3]
RP   FUNCTION.
RX   MEDLINE=96249383; PubMed=8647433;
```

```
RA   Zuo P., Maniatis T.;
RT   "The splicing factor U2AF35 mediates critical protein-protein
RT   interactions in constitutive and enhancer-dependent splicing.";
RL   Genes Dev. 10:1356-1368(1996).
CC   -!- FUNCTION: PLAYS A CRITICAL ROLE IN BOTH CONSTITUTIVE AND ENHANCER-
CC       DEPENDENT SPLICING BY MEDIATING ESSENTIAL PROTEIN-PROTEIN
CC       INTERACTIONS AND PROTEIN-RNA INTERACTIONS REQUIRED FOR ACCURATE 3'
CC       SPLICE SITE SELECTION. DIRECTLY MEDIATES INTERACTIONS BETWEEN
CC       U2AF65 (LARGE SUBUNIT) AND PROTEINS BOUND TO THE ENHANCERS AND
CC       THUS MAY FUNCTION AS A BRIDGE BETWEEN U2AF65 AND THE ENHANCER
CC       COMPLEX TO RECRUIT IT TO THE ADJACENT INTRON.
CC   -!- SUBUNIT: ASSOCIATES WITH A 65 kDa PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DOMAIN: THE AMINO-TERMINAL SR-RICH DOMAIN IS REQUIRED FOR
CC       INTERACTIONS WITH SR PROTEINS AND THE SPLICING REGULATORS TRA AND
CC       TRA2, AND THE CARBOXY-TERMINAL DOMAIN IS REQUIRED FOR FORMATION OF
CC       THE U2AF35/U2AF65 HETERODIMER.
CC   -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC   -!- SIMILARITY: CONTAINS 2 C3H1-TYPE ZINC FINGERS.
CC   -!- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M96982; AAA36619.1; -.
DR   EMBL; AP001748; BAA95534.1; -.
DR   PIR; S27969; S27969.
DR   PIR; A46179; A46179.
DR   MIM; 191317; -.
DR   InterPro; IPR000504; RRM.
DR   InterPro; IPR000571; Zf-CCCH.
DR   Pfam; PF00076; rrm; 1.
DR   Pfam; PF00642; zf-CCCH; 2.
DR   SMART; SM00360; RRM; 1.
DR   SMART; SM00356; ZnF_C3H1; 2.
DR   PROSITE; PS50102; RRM; 1.
DR   PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW   Nuclear protein; RNA-binding; mRNA splicing; Zinc-finger; Repeat.
FT   ZN_FING      13    41       C3H1-TYPE 1.
FT   DOMAIN       65    147      RNA-BINDING (RRM).
FT   ZN_FING     149    173      C3H1-TYPE 2.
FT   DOMAIN      178    240      ARG/GLY/SER-RICH (RS DOMAIN).
FT   DOMAIN      210    223      POLY-GLY.
SQ   SEQUENCE   240 AA;  27872 MW;  3DA130DCE0B953F6 CRC64;


  Query Match             68.9%;  Score 31;  DB 1;  Length 240;
  Best Local Similarity   71.4%;  Pred. No. 42;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 NATFYFK 8
             | :||||
Db      17 NCSFYFK 23


RESULT  11
U2AG_DROME
ID   U2AG_DROME     STANDARD;      PRT;   264 AA.
AC   Q94535; Q9VPN4;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Splicing factor U2af 38 kDa subunit (U2 auxiliary factor 38 kDa
DE   subunit) (U2 SNRNP auxiliary factor small subunit).
GN   U2AF38 OR CG3582.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Embryo;
RX   MEDLINE=96413646; PubMed=8816800;
RA   Rudner D.Z., Kanaar R., Breger K.S., Rio D.C.;
RT   "Mutations in the small subunit of the Drosophila U2AF splicing
RT   factor cause lethality and developmental defects.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:10333-10337(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: NECESSARY FOR THE SPLICING OF PRE-MRNA. BINDS TO THE
CC       POLYPYRIMIDINE TRACT OF INTRONS EARLY DURING SPLICEOSOME ASSEMBLY
CC       (BY SIMILARITY).
CC   -!- SUBUNIT: ASSOCIATES WITH A 65 kDa PROTEIN (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC   -!- SIMILARITY: CONTAINS 2 C3H1-TYPE ZINC FINGERS.
CC   -!- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U67066; AAB17271.1; -.
DR   EMBL; AE003590; AAF51512.1; -.
DR   FlyBase; FBgn0017457; U2af38.
DR   InterPro; IPR000504; RRM.
```

```
DR   InterPro; IPR003954; RRM_1.
DR   InterPro; IPR000571; Zf-CCCH.
DR   Pfam; PF00076; rrm; 1.
DR   Pfam; PF00642; zf-CCCH; 2.
DR   SMART; SM00361; RRM_1; 1.
DR   SMART; SM00356; ZnF_C3H1; 2.
DR   PROSITE; PS50102; RRM; 1.
DR   PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW   Nuclear protein; RNA-binding; mRNA splicing; Zinc-finger; Repeat.
FT   DOMAIN       44    149       RNA-BINDING (RRM).
FT   DOMAIN      180    213       ARG/SER-RICH (RS DOMAIN).
FT   DOMAIN      190    197       POLY-ARG.
FT   DOMAIN      252    262       POLY-GLY.
FT   CONFLICT     66     66       H -> D (IN REF. 1).
SQ   SEQUENCE   264 AA;  29877 MW;  577285FB66FDB2F5 CRC64;


  Query Match             68.9%;  Score 31;  DB 1;  Length 264;
  Best Local Similarity   71.4%;  Pred. No. 46;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 NATFYFK 8
          | :||||
Db     17 NCSFYFK 23


RESULT  12
ARGC_ECOLI
ID   ARGC_ECOLI     STANDARD;      PRT;   334 AA.
AC   P11446;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) (AGPR) (N-
DE   acetyl-glutamate semialdehyde dehydrogenase) (NAGSA dehydrogenase).
GN   ARGC OR B3958.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89121510; PubMed=2851495;
RA   Parsot C., Boyen A., Cohen G.N., Glansdorff N.;
RT   "Nucleotide sequence of Escherichia coli argB and argC genes:
RT   comparison of N-acetylglutamate kinase and
RT   N-acetylglutamate-gamma-semialdehyde dehydrogenase with homologous
RT   and analogous enzymes.";
RL   Gene 68:275-283(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=94089392; PubMed=8265357;
RA   Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA   Daniels D.L.;
RT   "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT   region from 89.2 to 92.8 minutes.";
RL   Nucleic Acids Res. 21:5408-5417(1993).
RN   [3]
RP   SEQUENCE OF 1-48 FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=83143275; PubMed=6761650;
RA   Piette J., Cunin R., Boyen A., Charlier D.R.M., Crabeel M.,
RA   van Vliet F., Glansdorff N., Squires C., Squires C.L.;
RT   "The regulatory region of the divergent argECBH operon in Escherichia
RT   coli K-12.";
RL   Nucleic Acids Res. 10:8031-8048(1982).
RN   [4]
RP   SEQUENCE OF 1-19 FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=92202162; PubMed=1551850;
RA   Meinnel T., Schmitt E., Mechulam Y., Blanquet S.;
RT   "Structural and biochemical characterization of the Escherichia coli
RT   argE gene product.";
RL   J. Bacteriol. 174:2323-2331(1992).
CC   -!- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+)
CC       + phosphate = N-acetyl-5-glutamyl phosphate + NADPH.
CC   -!- PATHWAY: THIRD STEP IN ARGININE BIOSYNTHESIS.
CC   -!- SIMILARITY: BELONGS TO THE NAGSA DEHYDROGENASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21446; AAA23477.1; -.
DR   EMBL; J01587; AAB59146.1; -.
DR   EMBL; X55417; -; NOT_ANNOTATED_CDS.
DR   EMBL; U00006; AAC43064.1; -.
DR   EMBL; AE000470; AAC76940.1; -.
DR   PIR; JT0332; RDECEP.
DR   EcoGene; EG10065; argC.
DR   InterPro; IPR000706; AGPR_act_site.
DR   InterPro; IPR000534; Semialdh_dh.
DR   Pfam; PF01118; Semialdhyde_dh; 1.
DR   Pfam; PF02774; Semialdhyde_dhC; 1.
DR   ProDom; PD003765; AGPR_act_site; 1.
DR   PROSITE; PS01224; ARGC; 1.
KW   Arginine biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT   ACT_SITE    154    154       BY SIMILARITY.
SQ   SEQUENCE   334 AA;  35952 MW;  67AC195ECE1C4789 CRC64;


  Query Match             68.9%;  Score 31;  DB 1;  Length 334;
  Best Local Similarity   75.0%;  Pred. No. 58;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 NNATFYFK 8
          |:|||| |
Db    109 NDATFYEK 116


RESULT  13
OMPU_VIBCH
ID   OMPU_VIBCH     STANDARD;      PRT;   341 AA.
AC   P97085; Q9L5A3; Q9KU90;
DT   15-DEC-1998 (Rel. 37, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Outer membrane protein U precursor (Porin ompU).
GN   OMPU OR VC0633.
OS   Vibrio cholerae.
OC   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX   NCBI_TaxID=666;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CLASSICAL OGAWA 395 / ATCC 39541 / SEROTYPE O1;
RX   MEDLINE=97101069; PubMed=8945596;
RA   Sperandio V., Bailey C.C., Giron J.A., Dirita V.J., Silveira W.D.,
RA   Vettore A.L., Kaper J.B.;
RT   "Cloning and characterization of the gene encoding the OmpU outer
RT   membrane protein of Vibrio cholerae.";
RL   Infect. Immun. 64:5406-5409(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EL TOR;
RA   Yin Y., Zhang J.Z.;
RT   "Cloning and expression of ompU of Vibrio cholerae and its
RT   antigenicity analysis.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EL TOR N16961 / SEROTYPE O1;
```

```
RX   MEDLINE=20406833; PubMed=10952301;
RA   Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA   Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA   Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA   Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA   McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA   Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA   Fraser C.M.;
RT   "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT   cholerae.";
RL   Nature 406:477-483(2000).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U73751; AAB48973.1; -.
DR   EMBL; AF253529; AAF64526.1; -.
DR   EMBL; AE004149; AAF93799.1; ALT_INIT.
DR   TIGR; VC0633; -.
DR   InterPro; IPR001702; Gram_neg_porin.
DR   InterPro; IPR003229; OMP_2.
DR   Pfam; PF00267; Gram-ve_porins; 1.
DR   ProDom; PD000808; OMP_2; 1.
KW   Transmembrane; Porin; Signal; Outer membrane; Complete proteome.
FT   SIGNAL       1     21       POTENTIAL.
FT   CHAIN       22    341       OUTER MEMBRANE PROTEIN U.
FT   CONFLICT   278    278       F -> I (IN REF. 2).
FT   CONFLICT   290    290       E -> K (IN REF. 1).
FT   CONFLICT   324    325       VG -> AS (IN REF. 1).
SQ   SEQUENCE   341 AA;  36645 MW;  CECB39070E441732 CRC64;


  Query Match             68.9%;  Score 31;  DB 1;  Length 341;
  Best Local Similarity   71.4%;  Pred. No. 59;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 NATFYFK 8
           :||:|||
Db     299 DATYYFK 305


RESULT  14
KNL2_BOVIN
ID   KNL2_BOVIN     STANDARD;      PRT;   434 AA.
AC   P01047;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Kininogen, LMW II precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83117859; PubMed=6572010;
RA   Nawa H., Kitamura N., Hirose T., Asai M., Inayama S., Nakanishi S.;
RT   "Primary structures of bovine liver low molecular weight kininogen
RT   precursors and their two mRNAs.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:90-94(1983).
RN   [2]
RP   SEQUENCE OF 19-376.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (3) THE
CC       ACTIVE PEPTIDE KALLIDIN THAT IS RELEASED FROM LMW-KININOGEN SHOWS
CC       A VARIETY OF PHYSIOLOGICAL EFFECTS: (3A) INFLUENCE IN SMOOTH
CC       MUSCLE CONTRACTION, (3B) INDUCTION OF HYPOTENSION, (3C)
CC       NATRIURESIS AND DIURESIS (KIDNEY).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS: HMW II AND LMW II KININOGEN PRECURSORS ARE
CC       PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
CC       SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
CC       TO RESIDUE 398.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- MISCELLANEOUS: LMW-KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT
CC       INVOLVED IN BLOOD CLOTTING.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00427; CAA23710.1; -.
DR   PIR; A01284; KGBOL2.
DR   HSSP; P01038; 1A90.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   Pfam; PF00031; cystatin; 3.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Signal.
FT   SIGNAL       1     18
FT   CHAIN       19    434       KININOGEN, LMW II.
FT   CHAIN       19    376       HEAVY CHAIN.
FT   PEPTIDE    378    386       BRADYKININ.
FT   CHAIN      387    434       LIGHT CHAIN.
FT   DOMAIN      19    135       CYSTATIN-LIKE 1.
FT   DOMAIN     136    256       CYSTATIN-LIKE 2.
FT   DOMAIN     257    376       CYSTATIN-LIKE 3.
FT   MOD_RES     19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD    87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   136    136       O-LINKED (PARTIAL).
FT   CARBOHYD   168    168       N-LINKED (GLCNAC. . .); OR 169.
FT   CARBOHYD   197    197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD   204    204       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   280    280       N-LINKED (GLCNAC. . .).
FT   DISULFID    27    404       INTERCHAIN.
FT   DISULFID    82     93
FT   DISULFID   106    125
FT   DISULFID   141    144
FT   DISULFID   205    217
FT   DISULFID   228    247
FT   DISULFID   261    264
FT   DISULFID   325    337
FT   DISULFID   348    367
SQ   SEQUENCE   434 AA;  48148 MW;  73A7079DE3E03430 CRC64;


  Query Match             68.9%;  Score 31;  DB 1;  Length 434;
  Best Local Similarity   62.5%;  Pred. No. 76;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 NNATFYFK 8
           :: |||||
```

```
Db    290 HDGTFYFK 297


RESULT 15
PAFA_HUMAN
ID   PAFA_HUMAN     STANDARD;      PRT;   441 AA.
AC   Q13093; Q15692;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Platelet-activating factor acetylhydrolase precursor (EC 3.1.1.47)
DE   (PAF acetylhydrolase) (PAF 2-acylhydrolase) (LDL-associated
DE   phospholipase A2) (LDL-PLA(2)) (2-acetyl-1-alkylglycerophosphocholine
DE   esterase) (1-alkyl-2-acetylglycerophosphocholine esterase).
GN   PLA2G7 OR PAFAH.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 42-57.
RC   TISSUE=Myeloid;
RX   MEDLINE=95214779; PubMed=7700381;
RA   Tjoelker L.W., Wilder C., Eberhardt C., Stafforini D.M., Dietsch G.,
RA   Schimpf B., Hooper S., le Trong H., Cousens L.S., Zimmerman G.A.,
RA   Yamada Y., McIntyre T.M., Prescott S.M., Gray P.W.;
RT   "Anti-inflammatory properties of a platelet-activating factor
RT   acetylhydrolase.";
RL   Nature 374:549-553(1995).
RN   [2]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Lymphoma;
RX   MEDLINE=96197208; PubMed=8624782;
RA   Tew D.G., Southan C., Rice S.Q.J., Lawrence M.P., Li H., Boyd H.F.,
RA   Moores K., Gloger I.S., Macphee C.H.;
RT   "Purification, properties, sequencing, and cloning of a lipoprotein-
RT   associated, serine-dependent phospholipase involved in the oxidative
RT   modification of low-density lipoproteins.";
RL   Arterioscler. Thromb. Vasc. Biol. 16:591-599(1996).
RN   [3]
RP   MUTAGENESIS.
RX   MEDLINE=96029630; PubMed=7592717;
RA   Tjoelker L.W., Eberhardt C., Unger J., le Trong H.,
RA   Zimmerman G.A., McIntyre T.M., Stafforini D.M., Prescott S.M.,
RA   Gray P.W.;
RT   "Plasma platelet-activating factor acetylhydrolase is a secreted
RT   phospholipase A2 with a catalytic triad.";
RL   J. Biol. Chem. 270:25481-25487(1995).
RN   [4]
RP   VARIANT PHE-279.
RX   MEDLINE=96259525; PubMed=8675689;
RA   Stafforini D.M., Satoh K., Atkinson D.L., Tjoelker L.W.,
RA   Eberhardt C., Yoshida H., Imaizumi T., Takamatsu S., Zimmerman G.A.,
RA   McIntyre T.M., Gray P.W., Prescott S.M.;
RT   "Platelet-activating factor acetylhydrolase deficiency. A missense
RT   mutation near the active site of an anti-inflammatory
RT   phospholipase.";
RL   J. Clin. Invest. 97:2784-2791(1996).
RN   [5]
RP   VARIANT PLA2G7 DEFICIENCY ARG-281.
RX   MEDLINE=97396177; PubMed=9245731;
RA   Yamada Y., Yokota M.;
RT   "Loss of activity of plasma platelet-activating factor acetylhydrolase
RT   due to a novel Gln281-->Arg mutation.";
RL   Biochem. Biophys. Res. Commun. 236:772-775(1997).
RN   [6]
RP   VARIANT PHE-279.
RX   MEDLINE=98074100; PubMed=9412624;
RA   Hiramoto M., Yoshida H., Imaizumi T., Yoshimizu N., Satoh K.;
RT   "A mutation in plasma platelet-activating factor acetylhydrolase
RT   (Val279-->Phe) is a genetic risk factor for stroke.";
RL   Stroke 28:2417-2420(1997).
RN   [7]
RP   VARIANT PHE-279.
RX   MEDLINE=98132308; PubMed=9472966;
RA   Yamada Y., Ichihara S., Fujimura T., Yokota M.;
RT   "Identification of the G994--> T missense in exon 9 of the plasma
RT   platelet-activating factor acetylhydrolase gene as an independent
RT   risk factor for coronary artery disease in Japanese men.";
RL   Metabolism 47:177-181(1998).
RN   [8]
RP   VARIANT PHE-279.
RX   MEDLINE=98430412; PubMed=9759612;
RA   Yoshida H., Imaizumi T., Fujimoto K., Itaya H., Hiramoto M.,
RA   Yoshimizu N., Fukushi K., Satoh K.;
RT   "A mutation in plasma platelet-activating factor acetylhydrolase
RT   (Val279Phe) is a genetic risk factor for cerebral hemorrhage but not
RT   for hypertension.";
RL   Thromb. Haemost. 80:372-375(1998).
RN   [9]
RP   VARIANTS HIS-92; THR-198 AND ALA-379.
RX   MEDLINE=20311534; PubMed=10733466;
RA   Kruse S., Mao X.-Q., Heinzmann A., Blattmann S., Roberts M.H.,
RA   Braun S., Gao P.-S., Forster J., Kuehr J., Hopkin J.M., Shirakawa T.,
RA   Deichmann K.A.;
RT   "The Ile198Thr and Ala379Val variants of plasmatic PAF-acetylhydrolase
RT   impair catalytical activities and are associated with atopy and
RT   asthma.";
RL   Am. J. Hum. Genet. 66:1522-1530(2000).
CC   -!- FUNCTION: MODULATES THE ACTION OF PLATELET-ACTIVATING FACTOR (PAF)
CC       BY HYDROLYZING THE SN-2 ESTER BOND TO YIELD THE BIOLOGICALLY
CC       INACTIVE LYSO-PAF. HAS A SPECIFICITY FOR SUBSTRATES WITH A SHORT
CC       RESIDUE AT THE SN-2 POSITION. IT IS INACTIVE AGAINST LONG-CHAIN
CC       PHOSPHOLIPIDS.
CC   -!- CATALYTIC ACTIVITY: 2-acetyl-1-alkyl-sn-glycero-3-phosphocholine +
CC       H(2)O = 1-alkyl-sn-glycero-3-phosphocholine + acetate.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- POLYMORPHISM: THE POLYMORPHIC ALLELES THR-198 AND VAL-379 ARE
CC       ASSOCIATED WITH ATOPY AND ASTHMA.
CC   -!- DISEASE: DEFECTS IN PLA2G7 ARE THE CAUSE OF PLA2G7 DEFICIENCY A
CC       TRAIT WHICH IS PRESENT IN 27% OF JAPANESE. IT COULD HAVE A
CC       SIGNIFICANT PHYSIOLOGIC EFFECT IN THE PRESENCE OF INFLAMMATORY
CC       BODILY RESPONSES.
CC   -!- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC       HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U20157; AAC50126.1; -.
DR   EMBL; U24577; AAB04170.1; -.
DR   MIM; 601690; -.
DR   InterPro; IPR000379; Est_lip_thioest_actsite.
DR   InterPro; IPR000734; Lipase.
DR   PROSITE; PS00120; LIPASE_SER; 1.
KW   Hydrolase; Lipid degradation; Glycoprotein; Signal; Polymorphism;
KW   Disease mutation.
FT   SIGNAL        1     21
FT   CHAIN        22    441       PLATELET-ACTIVATING FACTOR
FT                                 ACETYLHYDROLASE.
FT   ACT_SITE    273    273       CHARGE RELAY SYSTEM.
FT   ACT_SITE    296    296       CHARGE RELAY SYSTEM.
FT   ACT_SITE    351    351       CHARGE RELAY SYSTEM.
FT   CARBOHYD    423    423       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    433    433       N-LINKED (GLCNAC. . .).
FT   VARIANT      92     92       R -> H (COMMON POLYMORPHISM; IN
FT                                 DBSNP:1805017).
FT                                 /FTId=VAR_011583.
FT   VARIANT     198    198       I -> T (COMMON POLYMORPHISM; IN
FT                                 DBSNP:1805018).
```

```
FT                                     /FTId=VAR_011584.
FT   VARIANT      279    279           V -> F (IN PLA2G7 DEFICIENCY; LOSS OF
FT                                     FUNCTION; MORE COMMON AMONG JAPANESE THAN
FT                                     IN CAUCASIANS; RISK FACTOR FOR CORONARY
FT                                     ARTHERY DISEASE AND STROKE).
FT                                     /FTId=VAR_004268.
FT   VARIANT      281    281           Q -> R (IN PLA2G7 DEFICIENCY; LOSS OF
FT                                     FUNCTION).
FT                                     /FTId=VAR_011585.
FT   VARIANT      379    379           V -> A (COMMON POLYMORPHISM).
FT                                     /FTId=VAR_011586.
FT   MUTAGEN      108    108           S->A: ACTIVITY IS HIGHER THAN WILD TYPE.
FT   MUTAGEN      273    273           S->A: LOSS OF ACTIVITY.
FT   MUTAGEN      286    286           D->A: ALMOST NO ACTIVITY.
FT   MUTAGEN      286    286           D->N: DIMINISHED ACTIVITY.
FT   MUTAGEN      296    296           D->A: LOSS OF ACTIVITY.
FT   MUTAGEN      296    296           D->N: LOSS OF ACTIVITY.
FT   MUTAGEN      304    304           D->A: NO CHANGE IN ACTIVITY.
FT   MUTAGEN      338    338           D->A: ACTIVITY IS HIGHER THAN WILD TYPE.
FT   MUTAGEN      351    351           H->A: LOSS OF ACTIVITY.
SQ   SEQUENCE   441 AA;  50077 MW;  3BA9EEA9E8094A57 CRC64;


  Query Match             68.9%;  Score 31;  DB 1;  Length 441;
  Best Local Similarity   71.4%;  Pred. No. 77;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 NATFYFK 8
           :|||:|||
Db     185 SATYYFK 191


Search completed: July  1, 2002, 16:30:12
Job time: 621 sec
```

GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:18:06 ; Search time 75.26 Seconds
(without alignments)
18.389 Million cell updates/sec

Title: US-09-461-061A-1
Perfect score: 45
Sequence: 1 NNATFYFK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 36 | 80.0 | 858 | 5 | Q21853 | Q21853 caenorhabdi |
| 2 | 35 | 77.8 | 182 | 5 | Q21968 | Q21968 caenorhabdi |
| 3 | 34 | 75.6 | 71 | 16 | Q9KDI7 | Q9kdi7 bacillus ha |
| 4 | 34 | 75.6 | 115 | 15 | Q9YXT3 | Q9yxt3 human immun |
| 5 | 34 | 75.6 | 173 | 10 | Q9ZRF8 | Q9zrf8 oryza sativ |
| 6 | 34 | 75.6 | 181 | 10 | Q9M625 | Q9m625 hordeum vul |
| 7 | 34 | 75.6 | 182 | 10 | P93615 | P93615 triticum ae |
| 8 | 34 | 75.6 | 273 | 16 | Q9CD90 | Q9cd90 mycobacteri |
| 9 | 34 | 75.6 | 319 | 16 | Q9CJ05 | Q9cj05 lactococcus |
| 10 | 34 | 75.6 | 326 | 12 | Q66771 | Q66771 equine rota |
| 11 | 34 | 75.6 | 326 | 12 | Q66772 | Q66772 equine rota |
| 12 | 34 | 75.6 | 326 | 12 | Q91E87 | Q91e87 human rotav |
| 13 | 34 | 75.6 | 357 | 13 | Q91655 | Q91655 xenopus lae |
| 14 | 34 | 75.6 | 440 | 8 | Q9ZY23 | Q9zy23 pedinomonas |
| 15 | 34 | 75.6 | 497 | 10 | Q9FVS9 | Q9fvs9 arabidopsis |
| 16 | 34 | 75.6 | 602 | 2 | Q9AFA2 | Q9afa2 staphylococ |
| 17 | 34 | 75.6 | 630 | 5 | P91391 | P91391 caenorhabdi |
| 18 | 34 | 75.6 | 758 | 5 | O16610 | O16610 caenorhabdi |
| 19 | 34 | 75.6 | 949 | 16 | Q97DN1 | Q97dn1 clostridium |
| 20 | 33 | 73.3 | 251 | 10 | Q9LDL9 | Q9ldl9 arabidopsis |
| 21 | 33 | 73.3 | 303 | 5 | O44623 | O44623 caenorhabdi |
| 22 | 33 | 73.3 | 332 | 10 | O64597 | O64597 arabidopsis |
| 23 | 33 | 73.3 | 469 | 17 | O29037 | O29037 archaeoglob |
| 24 | 33 | 73.3 | 706 | 10 | Q9S7L6 | Q9s7l6 arabidopsis |
| 25 | 33 | 73.3 | 1346 | 5 | O45699 | O45699 caenorhabdi |
| 26 | 32 | 71.1 | 155 | 8 | O79080 | O79080 lipolexis g |
| 27 | 32 | 71.1 | 331 | 12 | Q91AP3 | Q91ap3 human rotav |
| 28 | 32 | 71.1 | 537 | 10 | Q9LH23 | Q9lh23 oryza sativ |
| 29 | 32 | 71.1 | 606 | 8 | Q9G1W4 | Q9g1w4 sus scrofa |
| 30 | 32 | 71.1 | 606 | 8 | Q9TDR1 | Q9tdr1 sus scrofa |
| 31 | 32 | 71.1 | 606 | 8 | O99997 | O99997 sus scrofa |
| 32 | 32 | 71.1 | 606 | 8 | Q9G7T6 | Q9g7t6 sus scrofa |
| 33 | 32 | 71.1 | 606 | 8 | Q9G7R8 | Q9g7r8 sus scrofa |
| 34 | 32 | 71.1 | 630 | 12 | Q9EML1 | Q9eml1 amsacta moo |
| 35 | 32 | 71.1 | 1096 | 4 | O94836 | O94836 homo sapien |
| 36 | 32 | 71.1 | 1191 | 5 | Q9VJN7 | Q9vjn7 drosophila |
| 37 | 32 | 71.1 | 1215 | 5 | Q22649 | Q22649 caenorhabdi |
| 38 | 32 | 71.1 | 1255 | 5 | Q9NK83 | Q9nk83 drosophila |
| 39 | 32 | 71.1 | 1560 | 5 | Q9GRX5 | Q9grx5 dictyosteli |
| 40 | 32 | 71.1 | 1561 | 5 | Q9U987 | Q9u987 dictyosteli |
| 41 | 32 | 71.1 | 1731 | 5 | Q95W43 | Q95w43 trypanosoma |
| 42 | 31 | 68.9 | 49 | 13 | Q98TR0 | Q98tr0 gallus gall |
| 43 | 31 | 68.9 | 110 | 11 | Q9JLF2 | Q9jlf2 rattus norv |
| 44 | 31 | 68.9 | 118 | 4 | Q9UI66 | Q9ui66 homo sapien |
| 45 | 31 | 68.9 | 121 | 15 | Q9Q7U5 | Q9q7u5 human immun |

ALIGNMENTS

RESULT 1
Q21853
ID Q21853 PRELIMINARY; PRT; 858 AA.
AC Q21853;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE R09A8.2 PROTEIN.
GN R09A8.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilkinson J.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z68009; CAA92004.1; -.
SQ SEQUENCE 858 AA; 96262 MW; 60E1070CA8F4D4C8 CRC64;

Query Match 80.0%; Score 36; DB 5; Length 858;
Best Local Similarity 75.0%; Pred. No. 89;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 NNATFYFK 8
            | :|||||
Db      49 NESTFYFK 56
```

RESULT 2
Q21968

```
ID   Q21968      PRELIMINARY;      PRT;   182 AA.
AC   Q21968;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   R12H7.3 PROTEIN.
GN   R12H7.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Coles L.;
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z50755; CAA90635.1; -.
DR   InterPro; IPR001232; Skp1.
DR   Pfam; PF01466; Skp1; 1.
DR   SMART; SM00512; Skp1; 1.
SQ   SEQUENCE   182 AA;  21178 MW;  4BE38A19C4FA8124 CRC64;


  Query Match            77.8%;  Score 35;  DB 5;  Length 182;
  Best Local Similarity  75.0%;  Pred. No. 31;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 NNATFYFK 8
           |||| :||
Db     161 NNATLFFK 168


RESULT   3
Q9KDI7
ID   Q9KDI7      PRELIMINARY;      PRT;    71 AA.
AC   Q9KDI7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   BH1226 PROTEIN.
GN   BH1226.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
DR   EMBL; AP001511; BAB04945.1; -.
KW   Complete proteome.
SQ   SEQUENCE   71 AA;  8026 MW;  311AC9AEB3C539D3 CRC64;


  Query Match            75.6%;  Score 34;  DB 16;  Length 71;
  Best Local Similarity  100.0%;  Pred. No. 19;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 NATFYF 7
           ||||||
Db      43 NATFYF 48


RESULT   4
Q9YXT3
ID   Q9YXT3      PRELIMINARY;      PRT;   115 AA.
AC   Q9YXT3;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ENVELOPE GLYCOPROTEIN C2V3 REGION (FRAGMENT).
GN   ENV.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RJ96BRP071;
RA   Tanuri A., Swanson P.A., Devare S.G., Berro O.J., Savedra A.,
RA   Costa L.J., Telles J.G., Brindeiro R., Schable C., Pieniazek D.,
RA   Rayfield M.;
RT   "HIV-1 subtypes among blood donors from Rio de Janeiro, Brazil.";
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF034019; AAC79271.1; -.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
KW   AIDS; Coat protein; Glycoprotein.
FT   NON_TER       1      1
FT   NON_TER     115    115
SQ   SEQUENCE   115 AA;  12755 MW;  6551E67B32DCF56C CRC64;


  Query Match            75.6%;  Score 34;  DB 15;  Length 115;
  Best Local Similarity  62.5%;  Pred. No. 31;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 NNATFYFK 8
           || ||:|:
Db      78 NNTTFFFR 85


RESULT   5
Q9ZRF8
ID   Q9ZRF8      PRELIMINARY;      PRT;   173 AA.
AC   Q9ZRF8;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYDROPHOBIC LEA-LIKE PROTEIN.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LOMELLO;
RA   Chen L.J., Chai Y.J., Chen P.W.;
RT   "A rice embryo-specific gene with high homology to soybean GmPM3 gene,
RT   a hydrophobic LEA protein gene.";
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U57639; AAD10377.1; -.
SQ   SEQUENCE   173 AA;  18287 MW;  63FA2F778BB7259D CRC64;


  Query Match            75.6%;  Score 34;  DB 10;  Length 173;
  Best Local Similarity  85.7%;  Pred. No. 46;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 NNATFYF 7
           | |||||
Db      47 NGATFYF 53
```

```
RESULT   6
Q9M625
ID   Q9M625      PRELIMINARY;      PRT;   181 AA.
AC   Q9M625;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PLASMA MEMBRANE ASSOCIATED PROTEIN.
GN   PM19.
OS   Hordeum vulgare (Barley).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Hordeum.
OX   NCBI_TaxID=4513;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Morris P.C., Ranford J.C.;
RT   "Hordeum vulgare hydrophobic embryo-associated protein PM19 mRNA.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF218627; AAF29532.1; -.
SQ   SEQUENCE   181 AA;  19036 MW;  849727F1123A4030 CRC64;


  Query Match             75.6%;  Score 34;  DB 10;  Length 181;
  Best Local Similarity   85.7%;  Pred. No. 48;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 NNATFYF 7
          | |||||
Db     47 NGATFYF 53


RESULT   7
P93615
ID   P93615      PRELIMINARY;      PRT;   182 AA.
AC   P93615;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ABA INDUCED PLASMA MEMBRANE PROTEIN PM 19.
GN   WTABAPM.
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Triticum.
OX   NCBI_TaxID=4565;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. CHIHOKU;
RX   MEDLINE=97393494; PubMed=9249988;
RA   Koike M., Takezawa D., Arakawa K., Yoshida S.;
RT   "Accumulation of 19-kDa plasma membrane polypeptide during induction
RT   of freezing tolerance in wheat suspension-cultured cells by abscisic
RT   acid.";
RL   Plant Cell Physiol. 38:707-716(1997).
DR   EMBL; U80037; AAB38504.1; -.
SQ   SEQUENCE   182 AA;  19009 MW;  799B9994AEA87DCB CRC64;


  Query Match             75.6%;  Score 34;  DB 10;  Length 182;
  Best Local Similarity   85.7%;  Pred. No. 48;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 NNATFYF 7
          | |||||
Db     47 NGATFYF 53


RESULT   8
Q9CD90
ID   Q9CD90      PRELIMINARY;      PRT;   273 AA.
AC   Q9CD90;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   HYPOTHETICAL PROTEIN ML0126.
GN   ML0126.
OS   Mycobacterium leprae.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1769;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TN;
RX   MEDLINE=21128732; PubMed=11234002;
RA   Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA   Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA   Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA   Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA   Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA   Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA   Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA   Barrell B.G.;
RT   "Massive gene decay in the leprosy bacillus.";
RL   Nature 409:1007-1011(2001).
DR   EMBL; AL583917; CAC29634.1; -.
DR   Leproma; ML0126; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   273 AA;  30831 MW;  D5B765A65EF39549 CRC64;


  Query Match             75.6%;  Score 34;  DB 16;  Length 273;
  Best Local Similarity   85.7%;  Pred. No. 71;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 NNATFYF 7
          ||| |||
Db     20 NNAQFYF 26


RESULT   9
Q9CJ05
ID   Q9CJ05      PRELIMINARY;      PRT;   319 AA.
AC   Q9CJ05;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RHAMNOSYLTRANSFERASE.
GN   RGPB.
OS   Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Lactococcus.
OX   NCBI_TaxID=1360;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IL1403;
RX   MEDLINE=21235186; PubMed=11337471;
RA   Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K.,
RA   Weissenbach J., Ehrlich S.D., Sorokin A.;
RT   "The complete genome sequence of the lactic acid bacterium Lactococcus
RT   lactis ssp. lactis IL1403.";
RL   Genome Res. 11:731-753(2001).
DR   EMBL; AE006258; AAK04299.1; -.
DR   HSSP; P39621; 1QG8.
DR   InterPro; IPR001173; Glycos_transf_2.
DR   Pfam; PF00535; Glycos_transf_2; 1.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   319 AA;  37340 MW;  D1DAC78AB0A950CA CRC64;


  Query Match             75.6%;  Score 34;  DB 16;  Length 319;
  Best Local Similarity   85.7%;  Pred. No. 82;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 NNATFYF 7
```

```
          ||| |||
Db     80 NNADFYF 86


RESULT  10
Q66771
ID   Q66771      PRELIMINARY;      PRT;   326 AA.
AC   Q66771;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   GLYCOPROTEIN VP7.
OS   Equine rotavirus.
OC   Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX   NCBI_TaxID=10937;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FR4;
RX   MEDLINE=95113988; PubMed=7814511;
RA   Ciarlet M., Reggeti F., Pina C.I., Liprandi F.;
RT   "Equine rotaviruses with G14 serotype specificity circulate among
RT   venezuelan horses.";
RL   J. Clin. Microbiol. 32:2609-2612(1994).
DR   EMBL; U05348; AAA81914.1; -.
DR   InterPro; IPR001963; VP7.
DR   Pfam; PF00434; VP7; 1.
SQ   SEQUENCE   326 AA;  37296 MW;  9FC1BD4F4CC76529 CRC64;


  Query Match             75.6%;  Score 34;  DB 12;  Length 326;
  Best Local Similarity   62.5%;  Pred. No. 84;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NNATFYFK 8
           |:|||::
Db     318 NSATFYYR 325


RESULT  11
Q66772
ID   Q66772      PRELIMINARY;      PRT;   326 AA.
AC   Q66772;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   GLYCOPROTEIN VP7.
OS   Equine rotavirus.
OC   Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX   NCBI_TaxID=10937;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FR8;
RX   MEDLINE=95113988; PubMed=7814511;
RA   Ciarlet M., Reggeti F., Pina C.I., Liprandi F.;
RT   "Equine rotaviruses with G14 serotype specificity circulate among
RT   venezuelan horses.";
RL   J. Clin. Microbiol. 32:2609-2612(1994).
DR   EMBL; U05349; AAA67342.1; -.
DR   InterPro; IPR001963; VP7.
DR   Pfam; PF00434; VP7; 1.
SQ   SEQUENCE   326 AA;  37352 MW;  9095E64B13933E29 CRC64;


  Query Match             75.6%;  Score 34;  DB 12;  Length 326;
  Best Local Similarity   62.5%;  Pred. No. 84;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NNATFYFK 8
           |:|||::
Db     318 NSATFYYR 325


RESULT  12
Q91E87
ID   Q91E87      PRELIMINARY;      PRT;   326 AA.
AC   Q91E87;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CAPSID PROTEIN.
GN   VP7.
OS   Human rotavirus.
OC   Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX   NCBI_TaxID=10941;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CC425, P3[9],G3;
RA   Griffin D.D., Nakagomi T., Hoshino Y., Nakagomi O., Kirkwood C.D.,
RA   Parashar U.D., Glass R.I., Gentsch J.R.;
RT   "Characterization of nontypeable rotavirus strains from the United
RT   States:Identification of a new rotavirus reassortant (P2A[6],G12) and
RT   rare P3[9] strains related to bovine rotaviruses.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ311738; CAC43312.1; -.
SQ   SEQUENCE   326 AA;  37119 MW;  8B97ED1DBBD1C981 CRC64;


  Query Match             75.6%;  Score 34;  DB 12;  Length 326;
  Best Local Similarity   62.5%;  Pred. No. 84;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NNATFYFK 8
           |:|||::
Db     318 NSATFYYR 325


RESULT  13
Q91655
ID   Q91655      PRELIMINARY;      PRT;   357 AA.
AC   Q91655;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENE 17 PROTEIN.
GN   GENE 17.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96312514; PubMed=8700860;
RA   Brown D.D., Wang Z., Furlow J.D., Kanamori A., Schwartzman R.A.,
RA   Remo B.F., Pinder A.;
RT   "The thyroid hormone-induced tail resorption program during Xenopus
RT   laevis metamorphosis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:1924-1929(1996).
DR   EMBL; U41860; AAC59876.1; -.
DR   InterPro; IPR001507; zona_pellucida.
DR   Pfam; PF00100; zona_pellucida; 1.
DR   SMART; SM00241; ZP; 1.
SQ   SEQUENCE   357 AA;  39090 MW;  5493352C8EEA21E6 CRC64;


  Query Match             75.6%;  Score 34;  DB 13;  Length 357;
  Best Local Similarity   85.7%;  Pred. No. 92;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 NNATFYF 7
           |||| ||
Db     141 NNATLYF 147


RESULT  14
Q9ZY23
```

```
ID   Q9ZY23      PRELIMINARY;      PRT;   440 AA.
AC   Q9ZY23;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NADH DEHYDROGENASE SUBUNIT 2 (EC 1.6.5.3).
GN   NAD2.
OS   Pedinomonas minor.
OG   Mitochondrion.
OC   Eukaryota; Viridiplantae; Chlorophyta; Pedinophyceae; Pedinomonadales;
OC   Pedinomonadaceae; Pedinomonas.
OX   NCBI_TaxID=3159;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=UTEX LB 1350;
RX   MEDLINE=99418884; PubMed=10488238;
RA   Turmel M., Lemieux C., Burger G., Lang B.F., Otis C., Plante I.,
RA   Gray M.W.;
RT   "The complete mitochondrial DNA sequences of Nephroselmis olivacea and
RT   Pedinomonas minor: two radically different evolutionary patterns
RT   within green algae.";
RL   Plant Cell 11:1717-1729(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=UTEX LB 1350;
RA   Burger G.;
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
DR   EMBL; AF116775; AAD19671.1; -.
DR   InterPro; IPR001750; Oxidored_q1.
DR   Pfam; PF00361; oxidored_q1; 1.
KW   Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ   SEQUENCE   440 AA;  51380 MW;  C6D6CBAD72549B4C CRC64;
```

```
  Query Match             75.6%;  Score 34;  DB 8;  Length 440;
  Best Local Similarity   85.7%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 NNATFYF 7
            ||| |||
Db    238 NNALFYF 244
```

```
RESULT  15
Q9FVS9
ID   Q9FVS9      PRELIMINARY;      PRT;   497 AA.
AC   Q9FVS9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CYTOCHROME P450, PUTATIVE.
GN   T8L23.21.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=21016719; PubMed=11130712;
RA   Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA   White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA   Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA   Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA   Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA   Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA   Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA   Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA   Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA   Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA   Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA   Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA   Sakano H., Salzerg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA   Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA   Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA   Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT   "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:816-820(2000).
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR   EMBL; AC079733; AAG50737.1; -.
DR   InterPro; IPR001128; Cyt_P450.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00385; P450.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Heme; Hypothetical protein; Monooxygenase; Oxidoreductase.
SQ   SEQUENCE   497 AA;  57696 MW;  C0A37BA59164AB78 CRC64;
```

```
  Query Match             75.6%;  Score 34;  DB 10;  Length 497;
  Best Local Similarity   85.7%;  Pred. No. 1.3e+02;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 NATFYFK 8
            | |||||
Db     64 NLTFYFK 70
```

```
Search completed: July  1, 2002, 16:29:44
Job time: 698 sec
```

GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:17:16 ; Search time 95.97 Seconds
(without alignments)
9.259 Million cell updates/sec

Title: US-09-461-061A-1
Perfect score: 45
Sequence: 1 NNATFYFK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 45 | 100.0 | 9 | 21 | AAB37455 | Human kininogen D3 |
| 2 | 45 | 100.0 | 10 | 21 | AAY95405 | Anti-angiogenic D3 |
| 3 | 45 | 100.0 | 16 | 21 | AAY95409 | Anti-angiogenic D3 |
| 4 | 45 | 100.0 | 16 | 21 | AAY95410 | Anti-angiogenic D3 |
| 5 | 45 | 100.0 | 32 | 21 | AAY95408 | Anti-angiogenic D3 |
| 6 | 45 | 100.0 | 117 | 14 | AAR33350 | Domaine 3, bradyki |
| 7 | 45 | 100.0 | 122 | 21 | AAB37447 | Human kininogen D3 |
| 8 | 45 | 100.0 | 123 | 21 | AAY95426 | Human high mol.wt. |
| 9 | 45 | 100.0 | 248 | 22 | ABG21102 | Novel human diagno |
| 10 | 45 | 100.0 | 369 | 22 | ABG21099 | Novel human diagno |
| 11 | 45 | 100.0 | 644 | 22 | ABG21101 | Novel human diagno |
| 12 | 39 | 86.7 | 26 | 18 | AAW54336 | Bradykinin analogo |
| 13 | 36 | 80.0 | 78 | 22 | AAO05080 | Human polypeptide |
| 14 | 34 | 75.6 | 42 | 22 | AAG74378 | Human colon cancer |
| 15 | 34 | 75.6 | 376 | 21 | AAB28404 | Staphylococcus hom |
| 16 | 34 | 75.6 | 541 | 22 | ABB52487 | Escherichia coli p |
| 17 | 34 | 75.6 | 602 | 20 | AAW93820 | Bacillus sp. GUS p |
| 18 | 34 | 75.6 | 602 | 20 | AAW93822 | Bacillus sp. GUS p |
| 19 | 34 | 75.6 | 602 | 20 | AAW93825 | Bacillus sp. codon |
| 20 | 34 | 75.6 | 602 | 20 | AAW93826 | Bacillus sp. GUS p |
| 21 | 34 | 75.6 | 602 | 21 | AAB28402 | Staphylococcus bet |
| 22 | 34 | 75.6 | 615 | 21 | AAB28408 | Codon-optimised St |
| 23 | 34 | 75.6 | 618 | 20 | AAW93821 | Bacillus sp. GUS p |
| 24 | 33 | 73.3 | 71 | 22 | AAM82831 | Human immune/haema |
| 25 | 33 | 73.3 | 173 | 21 | AAG44316 | Arabidopsis thalia |
| 26 | 33 | 73.3 | 251 | 21 | AAG44315 | Arabidopsis thalia |
| 27 | 33 | 73.3 | 259 | 21 | AAG44314 | Arabidopsis thalia |
| 28 | 32 | 71.1 | 1191 | 22 | ABB60775 | Drosophila melanog |
| 29 | 31 | 68.9 | 67 | 21 | AAB34049 | Human secreted pro |
| 30 | 31 | 68.9 | 77 | 22 | AAM84819 | Human immune/haema |
| 31 | 31 | 68.9 | 89 | 21 | AAG01238 | Human secreted pro |
| 32 | 31 | 68.9 | 121 | 21 | AAY81200 | Human mutant cysta |
| 33 | 31 | 68.9 | 128 | 21 | AAY81189 | Human mutant cysta |
| 34 | 31 | 68.9 | 149 | 20 | AAY36862 | Protein which is s |
| 35 | 31 | 68.9 | 161 | 22 | AAO12964 | Human polypeptide |
| 36 | 31 | 68.9 | 178 | 22 | ABG09760 | Novel human diagno |
| 37 | 31 | 68.9 | 193 | 16 | AAR71919 | Macaque platelet a |
| 38 | 31 | 68.9 | 193 | 18 | AAW26505 | Macaque platelet-a |
| 39 | 31 | 68.9 | 193 | 18 | AAW23800 | Macaque partial pl |
| 40 | 31 | 68.9 | 193 | 18 | AAW09812 | Partial macaque pl |
| 41 | 31 | 68.9 | 193 | 19 | AAW38365 | Monkey plasma plat |
| 42 | 31 | 68.9 | 193 | 20 | AAW96340 | Internal PAF-AH pe |
| 43 | 31 | 68.9 | 193 | 20 | AAW73365 | Macaque PAF-AH pro |
| 44 | 31 | 68.9 | 193 | 21 | AAB07996 | A macaque platelet |
| 45 | 31 | 68.9 | 193 | 21 | AAY88310 | Macaque PAF-AH ami |

ALIGNMENTS

RESULT 1
AAB37455
ID AAB37455 standard; peptide; 9 AA.
XX
AC AAB37455;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human kininogen D3 peptide fragment.
XX
KW Enzyme; legumain; endopeptidase; cystatin; human; kininogen.
XX
OS Homo sapiens.
XX
PN WO200064945-A1.
XX
PD 02-NOV-2000 bad date
XX
PF 20-APR-2000; 2000WO-GB01571.
XX
PR 22-APR-1999; 99GB-0009133.
XX
PA (BABR-) BABRAHAM INST.
XX
PI Abrahamson M, Barrett AJ;
XX
DR WPI; 2000-687316/67.
XX
PT Inhibition of mammalian legumain or legumain-related endopeptidase by
PT cystatin involves interaction with second papain-non-reactive site of
PT cystatin -
XX
PS Disclosure; Fig 4; 45pp; English.
XX

```
CC   The present invention relates to inhibition of the enzymatic activity of
CC   legumain or a legumain-related endopeptidase by cystatin. The inhibition
CC   involves an interaction between legumain and a papain-non-reactive site
CC   of cystatin. Legumain (EC 3.4.22.34) is a cysteine endopeptidase, and
CC   performs a protein-processing function. The present sequence is a peptide
CC   fragment of human kininogen D3, which was used in the present invention.
CC   Kininogen is a type 3 cystatin. The present sequence is thought to be
CC   involved in a legumain-inhibitory site.
XX
SQ   Sequence   9 AA;


  Query Match             100.0%;  Score 45;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NNATFYFK 8
          ||||||||
Db      1 nnatfyfk 8


RESULT   2
AAY95405
ID   AAY95405 standard; Peptide; 10 AA.
XX
AC   AAY95405;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Anti-angiogenic D3 peptide.
XX
KW   Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW   endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW   rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW   therapy; human; D3 peptide.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "Xaa represents 0 amino acids, or the
FT                          sequence of AAY95406 or its N-terminal
FT                          truncation fragment containing at least
FT                          1 amino acid"
FT   Peptide         2..9
FT                   /note= "corresponds to residues Asn(275)-Lys(282)
FT                          of HK domain 3"
FT   Misc-difference 10
FT                   /note= "Xaa represents 0 amino acids, or the
FT                          sequence of AAY95407 or its C-terminal
FT                          truncation fragment containing at least
FT                          1 amino acid"
XX
PN   WO200035407-A2.
XX
PD   22-JUN-2000.                    bad date
XX
PF   02-DEC-1999;   99WO-US28465.
XX
PR   16-DEC-1998;   98US-0112427.
XX
PA   (UTEM ) UNIV TEMPLE.
PA   (MCCR/) MCCRAE R K.
XX
PI   McCrae RK;
XX
DR   WPI; 2000-442247/38.
XX
PT   Composition for inhibiting angiogenesis and endothelial cell
PT   proliferation, inducing endothelial cell apoptosis and treating cancer,
PT   rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT   3 analog  -
XX
PS   Claim 1; Page 25; 44pp; English.
XX
CC   The present sequence is that of a D3 peptide derived from high
CC   mol.wt. kininogen (HK) domain 3 (see AAY95426).  The D3 peptide, which
CC   may optionally include N-terminal and/or C-terminal protecting
CC   groups, inhibits endothelial cell proliferation and thus possesses
CC   anti-angiogenic activity.  It is an example of peptides of the
CC   invention (see AAY95405-26) which are analogues of certain sites in
CC   the HK domain 3, in this case amino acids Asn275-Lys282.  The
CC   peptides inhibit endothelial cell proliferation and may also induce
CC   endothelial cell apoptosis.  Compositions including such peptides
CC   are used in claimed methods for inhibiting angiogenesis, inhibiting
CC   endothelial cell proliferation, and inducing endothelial cell
CC   apoptosis.  Cancer, rheumatoid arthritis, and ocular disorders
CC   characterized by undesired vascularization of the retina are treated.
XX
SQ   Sequence   10 AA;


  Query Match             100.0%;  Score 45;  DB 21;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.051;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NNATFYFK 8
          ||||||||
Db      2 nnatfyfk 9


RESULT   3
AAY95409
ID   AAY95409 standard; Peptide; 16 AA.
XX
AC   AAY95409;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Anti-angiogenic D3 peptide.
XX
KW   Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW   endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW   rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW   therapy; human; D3 peptide.
XX
OS   Homo sapiens.
XX
PN   WO200035407-A2.
XX
PD   22-JUN-2000.
XX
PF   02-DEC-1999;   99WO-US28465.
XX
PR   16-DEC-1998;   98US-0112427.
XX
PA   (UTEM ) UNIV TEMPLE.
PA   (MCCR/) MCCRAE R K.
XX
PI   McCrae RK;
XX
DR   WPI; 2000-442247/38.
XX
PT   Composition for inhibiting angiogenesis and endothelial cell
PT   proliferation, inducing endothelial cell apoptosis and treating cancer,
PT   rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT   3 analog  -
XX
PS   Claim 6; Page 26; 44pp; English.
XX
CC   The present sequence is that of a D3 peptide derived from human
CC   high mol.wt. kininogen (HK) domain 3 (see AAY95426).  The D3 peptide
CC   inhibits endothelial cell proliferation and thus possesses
CC   anti-angiogenic activity.  It is an example of D3 peptides of the
CC   invention (see AAY95405-26) that are analogues of certain sites in
CC   the HK domain 3, in this case amino acid residues Asn275-Lys282.
```

CC The peptides inhibit endothelial cell proliferation and may also
CC induce endothelial cell apoptosis. Compositions including the
CC peptides are used in claimed methods for inhibiting angiogenesis,
CC inhibiting endothelial cell proliferation, and inducing endothelial
CC cell apoptosis. Cancer, rheumatoid arthritis, and ocular disorders
CC characterized by undesired vascularization of the retina are treated.
CC The IC50 value for the present peptide was less than 0.8 uM for
CC inhibition of fibroblast growth factor-induced HUVEC cell
CC proliferation.
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 45; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
||||||||
Db 9 nnatfyfk 16

RESULT 4
AAY95410
ID AAY95410 standard; Peptide; 16 AA.
XX
AC AAY95410;
XX
DT 25-SEP-2000 (first entry)
XX
DE Anti-angiogenic D3 peptide.
XX
KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW therapy; human; D3 peptide.
XX
OS Homo sapiens.
XX
PN WO200035407-A2.
XX
PD 22-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28465.
XX
PR 16-DEC-1998; 98US-0112427.
XX
PA (UTEM ) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI McCrae RK;
XX
DR WPI; 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,
PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT 3 analog -
XX
PS Claim 7; Page 26; 44pp; English.
XX
CC The present sequence is that of a D3 peptide derived from human
CC high mol.wt. kininogen (HK) domain 3 (see AAY95426). The D3 peptide
CC inhibits endothelial cell proliferation and thus possesses
CC anti-angiogenic activity. It is an example of D3 peptides of the
CC invention (see AAY95405-26) that are analogues of certain sites in
CC the HK domain 3, in this case amino acid residues Asn275-Lys282.
CC The peptides inhibit endothelial cell proliferation and may also
CC induce endothelial cell apoptosis. Compositions including the
CC peptides are used in claimed methods for inhibiting angiogenesis,
CC inhibiting endothelial cell proliferation, and inducing endothelial
CC cell apoptosis. Cancer, rheumatoid arthritis, and ocular disorders
CC characterized by undesired vascularization of the retina are treated.
CC The IC50 value for the present peptide was less than 0.8 uM for
CC inhibition of fibroblast growth factor-induced HUVEC cell
CC proliferation.
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 45; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
||||||||
Db 1 nnatfyfk 8

RESULT 5
AAY95408
ID AAY95408 standard; Peptide; 32 AA.
XX
AC AAY95408;
XX
DT 25-SEP-2000 (first entry)
XX
DE Anti-angiogenic D3 peptide.
XX
KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW therapy; human; D3 peptide.
XX
OS Homo sapiens.
XX
PN WO200035407-A2.
XX
PD 22-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28465.
XX
PR 16-DEC-1998; 98US-0112427.
XX
PA (UTEM ) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI McCrae RK;
XX
DR WPI; 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,
PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT 3 analog -
XX
PS Claim 4; Page 26; 44pp; English.
XX
CC The present sequence is that of a D3 peptide derived from human
CC high mol.wt. kininogen (HK) domain 3 (see AAY95426). The D3 peptide
CC inhibits endothelial cell proliferation and thus possesses
CC anti-angiogenic activity. It is an example of D3 peptides of the
CC invention (see AAY95405-26) that are analogues of certain sites in
CC the HK domain 3, in this case amino acid residues Asn275-Lys282.
CC The peptides inhibit endothelial cell proliferation and may also
CC induce endothelial cell apoptosis. Compositions including the
CC peptides are used in claimed methods for inhibiting angiogenesis,
CC inhibiting endothelial cell proliferation, and inducing endothelial
CC cell apoptosis. Cancer, rheumatoid arthritis, and ocular disorders
CC characterized by undesired vascularization of the retina are treated.
XX
SQ Sequence 32 AA;

Query Match 100.0%; Score 45; DB 21; Length 32;
Best Local Similarity 100.0%; Pred. No. 0.17;

```
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NNATFYFK 8
           ||||||||
Db      13 nnatfyfk 20


RESULT   6
AAR33350
ID   AAR33350 standard; protein; 117 AA.
XX
AC   AAR33350;
XX
DT   01-JUL-1993  (first entry)
XX
DE   Domaine 3, bradykinin release activating peptide.
XX
KW   Domain 3; human; kininogen; heavy chain; low molecular weight; plasma;
KW   trypsin; platelet; activation; granule contents; hemostasis; thrombin;
KW   tissue plasminogen activator;thrombosis; inflammatory response;
KW   endothelial cell; von Willebrand factor;
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..18
FT                   /note= "Leader peptide"
FT   Protein         19..117
FT                   /note= "Mature protein"
XX
PN   WO9303748-A.
XX
PD   04-MAR-1993.
XX
PF   13-AUG-1992;   92WO-US06809.
XX
PR   13-AUG-1991;   91US-0744545.
XX
PA   (UTEM ) UNIV TEMPLE.
XX
PI   Jiang Y,  Schmaier AB;
XX
DR   WPI; 1993-093714/11.
XX
PT   Use of trypsin-cleavage fragment of human kininogen - for
PT   increasing vascular bradykinin release, for lowering blood
PT   pressure and treating hypertension
XX
PS   Disclosure; Fig 1; 46pp; English.
XX
CC   The sequence given represents domain 3, amino acids 246-362, of
CC   the human kininogen heavy chain.  Domain 3 was isolated from low
CC   molecular weight kininogen, derived from human plasma, by cleavage
CC   with trypsin.  Domain 3 peptide inhibits platelet activation causing
CC   a marked decrease in the platelets ability to aggregate and secrete
CC   their granule contents.  The granule contents comprise proteins which
CC   participate in hemostasis, thrombosis and the inflammatory response.
CC   Domain 3 also inhibits endothelial cell activation shown by a decrease
CC   in secretion of endothelial cell contents such as tissue plasminogen
CC   activator and von Willebrand factor.  Domain 3 functions to inhibit
CC   cell activation by blocking thrombin binding to its target cells, the
CC   peptide is a selective inhibitor of thrombin-induced platlet
CC   activation.
XX
SQ   Sequence   117 AA;


  Query Match             100.0%;  Score 45;  DB 14;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 0.64;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NNATFYFK 8
           ||||||||
Db      30 nnatfyfk 37


RESULT   7
AAB37447
ID   AAB37447 standard; protein; 122 AA.
XX
AC   AAB37447;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human kininogen D3.
XX
KW   Enzyme; legumain; endopeptidase; cystatin; human; kininogen.
XX
OS   Homo sapiens.
XX
PN   WO200064945-A1.
XX
PD   02-NOV-2000.
XX
PF   20-APR-2000; 2000WO-GB01571.
XX
PR   22-APR-1999;   99GB-0009133.
XX
PA   (BABR-) BABRAHAM INST.
XX
PI   Abrahamson M,  Barrett AJ;
XX
DR   WPI; 2000-687316/67.
XX
PT   Inhibition of mammalian legumain or legumain-related endopeptidase by
PT   cystatin involves interaction with second papain-non-reactive site of
PT   cystatin -
XX
PS   Disclosure; Fig 4; 45pp; English.
XX
CC   The present invention relates to inhibition of the enzymatic activity of
CC   legumain or a legumain-related endopeptidase by cystatin. The inhibition
CC   involves an interaction between legumain and a papain-non-reactive site
CC   of cystatin. Legumain (EC 3.4.22.34) is a cysteine endopeptidase, and
CC   performs a protein-processing function. The present sequence is human
CC   kininogen D3, which was used in the present invention. Kininogen is a
CC   type 3 cystatin.
XX
SQ   Sequence   122 AA;


  Query Match             100.0%;  Score 45;  DB 21;  Length 122;
  Best Local Similarity   100.0%;  Pred. No. 0.67;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NNATFYFK 8
           ||||||||
Db      35 nnatfyfk 42


RESULT   8
AAY95426
ID   AAY95426 standard; Peptide; 123 AA.
XX
AC   AAY95426;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Human high mol.wt. kininogen domain 3.
XX
KW   Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW   endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW   rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW   therapy; human; D3 peptide.
XX
OS   Homo sapiens.
```

XX
PN WO200035407-A2.
XX
PD 22-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28465.
XX
PR 16-DEC-1998; 98US-0112427.
XX
PA (UTEM ) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI McCrae RK;
XX
DR WPI; 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,
PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT 3 analog -
XX
PS Disclosure; Page 4; 44pp; English.
XX
CC The present sequence is that of domain 3 of human high mol.wt.
CC kininogen (HK). The invention provides peptides (see AAY95405-24)
CC that are analogues of certain sites in the HK domain 3,
CC specifically Asn275-Lys282, Cys246-Cys249, Leu331-Tyr338 and
CC Tyr299-Ser314. The peptides, in which native Cys residues may be
CC replaced by Ala residues, inhibit endothelial cell proliferation
CC and may also induce endothelial cell apoptosis. Compositions
CC including the peptides are used in claimed methods for inhibiting
CC angiogenesis, inhibiting endothelial cell proliferation, and
CC inducing endothelial cell apoptosis. Cancer, rheumatoid arthritis,
CC and ocular disorders characterized by undesired vascularization of
CC the retina are treated.
XX
SQ Sequence 123 AA;

Query Match 100.0%; Score 45; DB 21; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
||||||||
Db 41 nnatfyfk 48

RESULT 9
ABG21102
ID ABG21102 standard; Protein; 248 AA.
XX
AC ABG21102;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #21093.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS85289.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 51461; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 248 AA;

Query Match 100.0%; Score 45; DB 22; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
||||||||
Db 90 nnatfyfk 97

RESULT 10
ABG21099
ID ABG21099 standard; Protein; 369 AA.
XX
AC ABG21099;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #21090.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX

PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS85286.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 51458; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 369 AA;

Query Match 100.0%; Score 45; DB 22; Length 369;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
||||||||
Db 333 nnatfyfk 340

RESULT 11
ABG21101
ID ABG21101 standard; Protein; 644 AA.
XX
AC ABG21101;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #21092.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS85288.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 51460; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 644 AA;

Query Match 100.0%; Score 45; DB 22; Length 644;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
||||||||
Db 293 nnatfyfk 300

RESULT 12
AAW54336
ID AAW54336 standard; peptide; 26 AA.
XX
AC AAW54336;
XX
DT 30-JUL-1998 (first entry)
XX
DE Bradykinin analogous peptide 19.
XX
KW Inhibition; thrombin-induced platelet; prevention; platelet aggregation;
KW ADP-induced activation.
XX
OS Synthetic.
XX
PN WO9641640-A1.
XX
PD 27-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US09940.
XX
PR 09-JUN-1995; 95US-0000096.
XX
PA (UNMI ) UNIV MICHIGAN.
XX
PI Hasan AAK, Schmaier AH;
XX
DR WPI; 1997-065304/06.

```
XX
PT   Inhibition of platelet activation and aggregation - by admin. of new
PT   or known bradykinin analogues
XX
PS   Disclosure; Page 44; 73pp; English.
XX
CC   Administration of a peptide or multimer related to bradykinin or other
CC   disclosed peptides and multimers can be used for the inhibition of
CC   thrombin-induced platelets or other cells.  They can also be used for
CC   preventing platelet aggregation, or inhibiting ADP-induced activation.
CC   This is useful to prevent arterial occlusions arising from coronary
CC   thrombosis and stroke.
XX
SQ   Sequence   26 AA;


  Query Match             86.7%;  Score 39;  DB 18;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 NATFYFK 8
          |||||||
Db      1 natfyfk 7


RESULT  13
AAO05080
ID   AAO05080 standard; Protein; 78 AA.
XX
AC   AAO05080;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 18972.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorders; arthritis; inflammation.
XX
OS   Homo sapiens.
XX
PN   WO200164835-A2.
XX
PD   07-SEP-2001.
XX
PF   26-FEB-2001; 2001WO-US04927.
XX
PR   28-FEB-2000; 2000US-0515126.
PR   18-MAY-2000; 2000US-0577409.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-514838/56.
DR   N-PSDB; AAI85011.
XX
PT   Isolated nucleic acids and polypeptides, useful for preventing
PT   diagnosing and treating e.g. leukaemia, inflammation and immune
PT   disorders -
XX
PS   Claim 20; SEQ ID NO 18972; 1399pp + Sequence Listing; English.
XX
CC   The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC   the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   78 AA;


  Query Match             80.0%;  Score 36;  DB 22;  Length 78;
  Best Local Similarity   85.7%;  Pred. No. 18;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 NNATFYF 7
            || ||||
Db     38 nnttfyf 44


RESULT  14
AAG74378
ID   AAG74378 standard; Protein; 42 AA.
XX
AC   AAG74378;
XX
DT   03-SEP-2001  (first entry)
XX
DE   Human colon cancer antigen protein SEQ ID NO:5142.
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   colorectal carcinoma.
XX
OS   Homo sapiens.
XX
PN   WO200122920-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26524.
XX
PR   29-SEP-1999;   99US-0157137.
PR   03-NOV-1999;   99US-0163280.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Barash SC,  Birse CE,  Rosen CA;
XX
DR   WPI; 2001-235357/24.
DR   N-PSDB; AAH33809.
XX
PT   Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT   useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS   Claim 11; Page 6843; 9803pp; English.
XX
CC   AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC   cancer-associated nucleic acid molecules (N) and proteins (P), where
CC   the proteins are collectively known as colon cancer antigens. The colon
CC   cancer antigens have cytostatic activity and can be used in gene
CC   therapy and vaccine production. N and P may be used in the prevention,
CC   diagnosis and treatment of diseases associated with inappropriate P
CC   expression. For example, N and P may be used to treat disorders
CC   associated with decreased expression by rectifying mutations or deletions
CC   in a patient's genome that affect the activity of P by expressing
CC   inactive proteins or to supplement the patients own production of P.
CC   Additionally, N may be used to produce the colon cancer-associated Ps,
CC   by inserting the nucleic acids into a host cell and culturing the cell
CC   to express the proteins. N and P can be used in the prevention, diagnosis
CC   and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC   and AAB77789 represent sequences used in the exemplification of the
CC   present invention.
CC   N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC   missing at time of publication, meaning no sequences are present for
CC   SEQ ID NO:1027 to 1052, 7921 and 7922.
```

XX
SQ Sequence 42 AA;

Query Match 75.6%; Score 34; DB 22; Length 42;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NNATFYF 7
| |||||
Db 22 nratfyf 28

RESULT 15
AAB28404
ID AAB28404 standard; Protein; 376 AA.
XX
AC AAB28404;
XX
DT 26-JAN-2001 (first entry)
XX
DE Staphylococcus homini beta-D-glucoronidase.
XX
KW Microbial; beta-glucoronidase; GUS; Enterobacter; Salmonella;
KW Pseudomonas; Staphylococcus; Thermotoga; transgenic plant; bioindicator;
KW transgenic insect; marker; glucuronide detoxification.
XX
OS Staphylococcus homini.
XX
PN WO200055333-A1.
XX
PD 21-SEP-2000.
XX
PF 16-MAR-2000; 2000WO-US07107.
XX
PR 17-MAR-1999; 99US-0270957.
XX
PA (CAMB-) CAMBIA BIOSYSTEMS LLC.
XX
PI Jefferson RA, Mayer JE;
XX
DR WPI; 2000-647075/62.
DR N-PSDB; AAA07936.
XX
PT Novel microbial beta-glucuronidase genes and gene products used as
PT reporter/effector molecule, as diagnostic tool, in positive selection,
PT to target molecules to specific cells and to detect and track linked
PT genes -
XX
PS Claim 3; Fig 5B; 116pp; English.
XX
CC The present sequence is a microbial beta-glucuronidase (GUS)
CC protein. GUS genes were obtained from six different genera:
CC Enterobacter/Salmonella, Pseudomonas, Salmonella, Staphylococcus and
CC Thermotoga. Microbial GUS can be used as a reporter/effector molecule for
CC transgenic constructions and in in vitro diagnostic applications. It may
CC also be used to generate sentinel plants that serve as bioindicators of
CC environmental status. It may be used to generate transgenic insects for
CC tracking insect populations or to facilitate the development of a
CC bioassay for compounds that affect molecules critical for insect
CC development (e.g. juvenile hormone). Secreted GUS may also serve as a
CC marker for beneficial fungi destined for release into the environment. In
CC animal systems, secreted GUS may be used to achieve extracellular
CC detoxification of glucuronides (e.g. toxin glucuronide) and to examine
CC conjugation patterns of glucuronides. Microbial GUS may also be used in
CC traditional medical diagnostic assays, for drug testing, pharmacokinetic
CC studies, bioavailability studies, diagnosis of diseases and syndromes,
CC following progression of disease or its response to therapy. Microbial
CC GUS has increased thermal stability, high turnover number and enzymatic
CC activity. It is highly specific for the substrate and water soluble, and
CC the substrates are stable.
XX
SQ Sequence 376 AA;

Query Match 75.6%; Score 34; DB 21; Length 376;
Best Local Similarity 75.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
|| ||||
Db 57 nnkpfyfk 64

Search completed: July 1, 2002, 16:19:43
Job time: 147 sec

GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:30:12 ; Search time 21.51 Seconds
(without alignments)
21.601 Million cell updates/sec

Title: US-09-461-061A-2
Perfect score: 60
Sequence: 1 TLTHTITKLNAE 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 60 | 100.0 | 644 | 1 | KNG_HUMAN | P01042 homo sapien |
| 2 | 42 | 70.0 | 436 | 1 | KNL1_BOVIN | P01046 bos taurus |
| 3 | 42 | 70.0 | 621 | 1 | KNH1_BOVIN | P01044 bos taurus |
| 4 | 41 | 68.3 | 434 | 1 | KNL2_BOVIN | P01047 bos taurus |
| 5 | 41 | 68.3 | 617 | 1 | VGF_RAT | P20156 rattus norv |
| 6 | 41 | 68.3 | 619 | 1 | KNH2_BOVIN | P01045 bos taurus |
| 7 | 37 | 61.7 | 1969 | 1 | MYSA_CAEEL | P12844 caenorhabdi |
| 8 | 36 | 60.0 | 761 | 1 | SM3D_CHICK | Q90663 gallus gall |
| 9 | 36 | 60.0 | 1231 | 1 | KF4A_MOUSE | P33174 mus musculu |
| 10 | 35 | 58.3 | 150 | 1 | YMEL_STRLN | P55049 streptomyce |
| 11 | 35 | 58.3 | 178 | 1 | HSLV_XYLFA | Q9pd95 xylella fas |
| 12 | 35 | 58.3 | 291 | 1 | COAT_LSV | P27335 lily sympto |
| 13 | 35 | 58.3 | 639 | 1 | KNG_RAT | P08934 rattus norv |
| 14 | 35 | 58.3 | 661 | 1 | KNG_MOUSE | O08677 mus musculu |
| 15 | 35 | 58.3 | 1211 | 1 | ATH1_YEAST | P48016 saccharomyc |
| 16 | 35 | 58.3 | 3010 | 1 | POLG_HCVBK | P26663 h genome po |
| 17 | 35 | 58.3 | 3010 | 1 | POLG_HCVJA | P26662 h genome po |
| 18 | 35 | 58.3 | 3010 | 1 | POLG_HCVJT | Q00269 h genome po |
| 19 | 35 | 58.3 | 3010 | 1 | POLG_HCVTW | P29846 h genome po |
| 20 | 35 | 58.3 | 3011 | 1 | POLG_HCVH | P27958 h genome po |
| 21 | 34 | 56.7 | 218 | 1 | TPMT_VIBCH | Q9ksn0 vibrio chol |
| 22 | 34 | 56.7 | 395 | 1 | EFTU_XYLFA | Q9p9q9 xylella fas |
| 23 | 34 | 56.7 | 452 | 1 | EFTU_BOVIN | P49410 bos taurus |
| 24 | 34 | 56.7 | 452 | 1 | EFTU_HUMAN | P49411 homo sapien |
| 25 | 34 | 56.7 | 565 | 1 | Y322_MYCPN | P75323 mycoplasma |
| 26 | 34 | 56.7 | 898 | 1 | KBF2_HUMAN | Q00653 homo sapien |
| 27 | 34 | 56.7 | 1028 | 1 | BGAL_ENTCL | Q47077 enterobacte |
| 28 | 34 | 56.7 | 1081 | 1 | UL52_HSVEB | P28962 equine herp |
| 29 | 34 | 56.7 | 3011 | 1 | POLG_HCV1 | P26664 h genome po |
| 30 | 34 | 56.7 | 3033 | 1 | POLG_HCVJ6 | P26660 h genome po |
| 31 | 34 | 56.7 | 3033 | 1 | POLG_HCVJ8 | P26661 h genome po |
| 32 | 33 | 55.0 | 228 | 1 | NUSG_MYCLE | Q9cbk0 mycobacteri |
| 33 | 33 | 55.0 | 234 | 1 | TNFA_HORSE | P29553 equus cabal |
| 34 | 33 | 55.0 | 238 | 1 | NUSG_MYCTU | P96930 mycobacteri |
| 35 | 33 | 55.0 | 249 | 1 | CYSH_SYNY3 | P72794 synechocyst |
| 36 | 33 | 55.0 | 265 | 1 | YFU2_BACST | Q04729 bacillus st |
| 37 | 33 | 55.0 | 389 | 1 | EFT3_STRRA | P29544 streptomyce |
| 38 | 33 | 55.0 | 392 | 1 | EFT3_STRCO | P40175 streptomyce |
| 39 | 33 | 55.0 | 446 | 1 | ENO_PLAFA | Q27727 plasmodium |
| 40 | 33 | 55.0 | 494 | 1 | ATF7_HUMAN | P17544 homo sapien |
| 41 | 33 | 55.0 | 599 | 1 | MX67_YEAST | Q99257 saccharomyc |
| 42 | 33 | 55.0 | 835 | 1 | FASD_ECOLI | P46000 escherichia |
| 43 | 33 | 55.0 | 1194 | 1 | MGR1_HUMAN | Q13255 homo sapien |
| 44 | 33 | 55.0 | 1199 | 1 | MGR1_RAT | P23385 rattus norv |
| 45 | 32 | 53.3 | 77 | 1 | PSK5_ARATH | Q9feb4 arabidopsis |

ALIGNMENTS

RESULT 1
KNG_HUMAN
ID KNG_HUMAN STANDARD; PRT; 644 AA.
AC P01042; P01043;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Kininogen precursor (Alpha-2-thiol proteinase inhibitor) [Contains:
DE Bradykinin].
GN KNG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC TISSUE=Liver;
RX MEDLINE=85234582; PubMed=2989293;
RA Takagaki Y., Kitamura N., Nakanishi S.;
RT "Cloning and sequence analysis of cDNAs for human high molecular
RT weight and low molecular weight prekininogens. Primary structures of
RT two human prekininogens.";
RL J. Biol. Chem. 260:8601-8609(1985).
RN [2]
RP GENE STRUCTURE.
RX MEDLINE=85234583; PubMed=2989294;
RA Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T.,
RA Nakanishi S.;
RT "Structural organization of the human kininogen gene and a model for
RT its evolution.";
RL J. Biol. Chem. 260:8610-8617(1985).
RN [3]
RP SEQUENCE OF 1-401 FROM N.A.
RX MEDLINE=85122621; PubMed=6441591;
RA Ohkubo I., Kurachi K., Takasawa T., Shiokawa H., Sasaki M.;
RT "Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and
RT its identity with low molecular weight kininogen.";
RL Biochemistry 23:5691-5697(1984).
RN [4]
RP SEQUENCE OF 379-644.
RX MEDLINE=86030270; PubMed=4054110;
RA Lottspeich F., Kellermann J., Henschen A., Foertsch B.,
RA Mueller-Esterl W.;
RT "The amino acid sequence of the light chain of human high-molecular-
RT mass kininogen.";
RL Eur. J. Biochem. 152:307-314(1985).
RN [5]
RP SEQUENCE OF 381-389.
RX MEDLINE=90255622; PubMed=4952632;
RA Pierce J.V.;
RT "Structural features of plasma kinins and kininogens.";
RL Fed. Proc. 27:52-57(1968).
RN [6]
RP DISULFIDE BONDS.
RA Sueyoshi T., Miyata T., Kato H., Iwanaga S.;
RT "Disulfide bonds in bovine HMW kininogens.";

```
RL   Seikagaku 56:808-808(1984).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC       HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
CC       FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-
CC       INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
CC       BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
CC       PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
CC       CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
CC       DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
CC       MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
CC       PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
CC       OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
CC       A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
CC       INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
CC       KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
CC       CLOTTING.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; K02566; AAA35497.1; -.
DR   EMBL; M11437; AAB59550.1; -.
DR   EMBL; M11438; AAB59550.1; JOINED.
DR   EMBL; M11521; AAB59550.1; JOINED.
DR   EMBL; M11522; AAB59550.1; JOINED.
DR   EMBL; M11523; AAB59550.1; JOINED.
DR   EMBL; M11524; AAB59550.1; JOINED.
DR   EMBL; M11525; AAB59550.1; JOINED.
DR   EMBL; M11526; AAB59550.1; JOINED.
DR   EMBL; M11527; AAB59550.1; JOINED.
DR   EMBL; M11528; AAB59550.1; JOINED.
DR   EMBL; M11437; AAB59551.1; -.
DR   EMBL; M11438; AAB59551.1; JOINED.
DR   EMBL; M11521; AAB59551.1; JOINED.
DR   EMBL; M11522; AAB59551.1; JOINED.
DR   EMBL; M11523; AAB59551.1; JOINED.
DR   EMBL; M11524; AAB59551.1; JOINED.
DR   EMBL; M11525; AAB59551.1; JOINED.
DR   EMBL; M11526; AAB59551.1; JOINED.
DR   EMBL; M11527; AAB59551.1; JOINED.
DR   EMBL; M11528; AAB59551.1; JOINED.
DR   PIR; A01279; KGHUH1.
DR   PIR; A25276; A25276.
DR   PIR; A01280; KGHUL1.
DR   PIR; B25276; B25276.
DR   PIR; S02482; S02482.
DR   SWISS-2DPAGE; P01042; HUMAN.
DR   MIM; 228960; -.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW   Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW   Alternative splicing.
FT   SIGNAL       1     18
FT   CHAIN        19    644       KININOGEN.
FT   CHAIN        19    380       KININOGEN HEAVY CHAIN.
FT   PEPTIDE     381    389       BRADYKININ.
FT   CHAIN       390    644       KININOGEN LIGHT CHAIN.
FT   DOMAIN       19    136       CYSTATIN-LIKE 1.
FT   DOMAIN      137    258       CYSTATIN-LIKE 2.
FT   DOMAIN      259    380       CYSTATIN-LIKE 3.
FT   DOMAIN      420    510       HIS-RICH (ASSOCIATED WITH CLOTTING
FT                                ACTIVITY).
FT   REPEAT      420    449
FT   REPEAT      450    479
FT   REPEAT      480    510
FT   MOD_RES      19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   DISULFID     28    614       INTERCHAIN.
FT   DISULFID     83     94
FT   DISULFID    107    126
FT   DISULFID    142    145
FT   DISULFID    206    218
FT   DISULFID    229    248
FT   DISULFID    264    267
FT   DISULFID    328    340
FT   DISULFID    351    370
FT   CARBOHYD     48     48       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    169    169       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    205    205       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    294    294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    401    401       O-LINKED.
FT   CARBOHYD    533    533       O-LINKED.
FT   CARBOHYD    542    542       O-LINKED.
FT   CARBOHYD    546    546       O-LINKED.
FT   CARBOHYD    557    557       O-LINKED.
FT   CARBOHYD    571    571       O-LINKED.
FT   CARBOHYD    577    577       O-LINKED.
FT   CARBOHYD    593    593       O-LINKED.
FT   CARBOHYD    628    628       O-LINKED.
FT   VARSPLIC    402    427       VSPPHTSMAPAQDEERDSGKEQGHTR -> SHLRSCEYKGR
FT                                PPKAGAEPASEREVS (IN ISOFORM LMW).
FT   VARSPLIC    428    644       MISSING (IN ISOFORM LMW).
FT   CONFLICT    593    593       T -> I (IN REF. 1).
SQ   SEQUENCE   644 AA;  71945 MW;  3132B4CBAF8FBB7E CRC64;


  Query Match            100.0%;  Score 60;  DB 1;  Length 644;
  Best Local Similarity  100.0%;  Pred. No. 0.00075;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAE 12
           ||||||||||||
Db     281 TLTHTITKLNAE 292


RESULT   2
KNL1_BOVIN
ID   KNL1_BOVIN     STANDARD;      PRT;   436 AA.
AC   P01046;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Kininogen, LMW I precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83117859; PubMed=6572010;
RA   Nawa H., Kitamura N., Hirose T., Asai M., Inayama S., Nakanishi S.;
RT   "Primary structures of bovine liver low molecular weight kininogen
RT   precursors and their two mRNAs.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:90-94(1983).
RN   [2]
```

```
RP   SEQUENCE OF 19-378.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (3) THE
CC       ACTIVE PEPTIDE KALLIDIN THAT IS RELEASED FROM LMW-KININOGEN SHOWS
CC       A VARIETY OF PHYSIOLOGICAL EFFECTS: (3A) INFLUENCE IN SMOOTH
CC       MUSCLE CONTRACTION, (3B) INDUCTION OF HYPOTENSION, (3C)
CC       NATRIURESIS AND DIURESIS (KIDNEY).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS: HMW I AND LMW I KININOGEN PRECURSORS ARE
CC       PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
CC       SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
CC       TO RESIDUE 400.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- MISCELLANEOUS: LMW-KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT
CC       INVOLVED IN BLOOD CLOTTING.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00426; CAA23709.1; -.
DR   EMBL; J00010; AAA30604.1; -.
DR   PIR; A01283; KGBOL1.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   Pfam; PF00031; cystatin; 3.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Signal.
FT   SIGNAL        1     18
FT   CHAIN        19    436       KININOGEN, LMW I.
FT   CHAIN        19    378       HEAVY CHAIN.
FT   PEPTIDE     380    388       BRADYKININ.
FT   CHAIN       389    436       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    257       CYSTATIN-LIKE 2.
FT   DOMAIN      258    378       CYSTATIN-LIKE 3.
FT   MOD_RES      19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    136    136       O-LINKED (PARTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .); OR 169.
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .).
FT   DISULFID     27    406       INTERCHAIN.
FT   DISULFID     82     93
FT   DISULFID    106    125
FT   DISULFID    141    144
FT   DISULFID    205    217
FT   DISULFID    228    247
FT   DISULFID    263    266
FT   DISULFID    327    339
FT   DISULFID    350    369
FT   CONFLICT    295    295       A -> T (IN REF. 1; CAA23709).
SQ   SEQUENCE   436 AA;  48427 MW;  F01F7EB6814BCE6C CRC64;


  Query Match             70.0%;  Score 42;  DB 1;  Length 436;
  Best Local Similarity   72.7%;  Pred. No. 1.4;

    Matches    8;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAE 12
           |:|:| |||||
Db     281 LSHSIAKLNAE 291


RESULT   3
KNH1_BOVIN
ID   KNH1_BOVIN     STANDARD;      PRT;   621 AA.
AC   P01044;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Kininogen, HMW I precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84014106; PubMed=6571699;
RA   Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
RT   "A single gene for bovine high molecular weight and low molecular
RT   weight kininogens.";
RL   Nature 305:545-549(1983).
RN   [2]
RP   SEQUENCE OF 19-378.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
RN   [3]
RP   SEQUENCE OF 378-393.
RX   MEDLINE=70180420; PubMed=4986212;
RA   Kato H., Nagasawa S., Suzuki T.;
RT   "Studies on the structure of bovine kininogen: cleavages of disulfide
RT   bonds and of methionyl bonds in kininogen-II.";
RL   J. Biochem. 67:313-323(1970).
RN   [4]
RP   SEQUENCE OF 458-498.
RX   MEDLINE=75170265; PubMed=1169237;
RA   Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
RT   "Studies on the primary structure of bovine high-molecular-weight
RT   kininogen. Amino acid sequence of a fragment ('histidine-rich
RT   peptide') released by plasma kallikrein.";
RL   J. Biochem. 77:55-68(1975).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC       HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
CC       TO FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
CC       PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE
CC       PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
CC       A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
CC       MUSCLE CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C)
CC       NATRIURESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
CC       (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
CC       IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
CC       RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
CC       (4F) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
CC       ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
CC       ACTION).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS: HMW I AND LMW I KININOGEN PRECURSORS ARE
CC       PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
CC       SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
CC       TO RESIDUE 400.
CC   -!- TISSUE SPECIFICITY: PLASMA.
```

```
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V01491; CAA24735.1; -.
DR   PIR; A01281; KGBOH1.
DR   PIR; A29559; A29559.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Blood coagulation;
KW   Inflammatory response; Signal.
FT   SIGNAL        1     18       PROBABLE.
FT   CHAIN        19    621       KININOGEN, HMW I.
FT   CHAIN        19    378       HEAVY CHAIN.
FT   PEPTIDE     380    388       BRADYKININ.
FT   CHAIN       389    621       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    257       CYSTATIN-LIKE 2.
FT   DOMAIN      258    378       CYSTATIN-LIKE 3.
FT   MOD_RES      19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    136    136       O-LINKED (PARTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .); OR 169.
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .).
FT   DISULFID     27    591       INTERCHAIN.
FT   DISULFID     82     93
FT   DISULFID    106    125
FT   DISULFID    141    144
FT   DISULFID    205    217
FT   DISULFID    228    247
FT   DISULFID    263    266
FT   DISULFID    327    339
FT   DISULFID    350    369
SQ   SEQUENCE   621 AA;  68890 MW;  D16850BEFE3C55CD CRC64;


  Query Match          70.0%;  Score 42;  DB 1;  Length 621;
  Best Local Similarity  72.7%;  Pred. No. 2;
  Matches    8;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAE 12
           |:|:| |||||
Db     281 LSHSIAKLNAE 291


RESULT   4
KNL2_BOVIN
ID   KNL2_BOVIN     STANDARD;      PRT;   434 AA.
AC   P01047;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Kininogen, LMW II precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83117859; PubMed=6572010;
RA   Nawa H., Kitamura N., Hirose T., Asai M., Inayama S., Nakanishi S.;
RT   "Primary structures of bovine liver low molecular weight kininogen
RT   precursors and their two mRNAs.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:90-94(1983).
RN   [2]
RP   SEQUENCE OF 19-376.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (3) THE
CC       ACTIVE PEPTIDE KALLIDIN THAT IS RELEASED FROM LMW-KININOGEN SHOWS
CC       A VARIETY OF PHYSIOLOGICAL EFFECTS: (3A) INFLUENCE IN SMOOTH
CC       MUSCLE CONTRACTION, (3B) INDUCTION OF HYPOTENSION, (3C)
CC       NATRIURESIS AND DIURESIS (KIDNEY).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS: HMW II AND LMW II KININOGEN PRECURSORS ARE
CC       PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
CC       SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
CC       TO RESIDUE 398.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- MISCELLANEOUS: LMW-KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT
CC       INVOLVED IN BLOOD CLOTTING.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00427; CAA23710.1; -.
DR   PIR; A01284; KGBOL2.
DR   HSSP; P01038; 1A90.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   Pfam; PF00031; cystatin; 3.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Signal.
FT   SIGNAL        1     18
FT   CHAIN        19    434       KININOGEN, LMW II.
FT   CHAIN        19    376       HEAVY CHAIN.
FT   PEPTIDE     378    386       BRADYKININ.
FT   CHAIN       387    434       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    256       CYSTATIN-LIKE 2.
FT   DOMAIN      257    376       CYSTATIN-LIKE 3.
FT   MOD_RES      19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    136    136       O-LINKED (PARTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .); OR 169.
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    280    280       N-LINKED (GLCNAC. . .).
FT   DISULFID     27    404       INTERCHAIN.
FT   DISULFID     82     93
FT   DISULFID    106    125
FT   DISULFID    141    144
FT   DISULFID    205    217
```

```
FT   DISULFID    228    247
FT   DISULFID    261    264
FT   DISULFID    325    337
FT   DISULFID    348    367
SQ   SEQUENCE   434 AA;  48148 MW;  73A7079DE3E03430 CRC64;


  Query Match             68.3%;  Score 41;  DB 1;  Length 434;
  Best Local Similarity   72.7%;  Pred. No. 2.1;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAE 12
           | |:| |||||
Db     279 LNHSIAKLNAE 289


RESULT   5
VGF_RAT
ID   VGF_RAT        STANDARD;      PRT;   617 AA.
AC   P20156;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   VGF protein precursor (VGF8A protein).
GN   VGF.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91203852; PubMed=2017159;
RA   Salton S.R.J., Fischberg D.J., Dong K.-W.;
RT   "Structure of the gene encoding VGF, a nervous system-specific mRNA
RT   that is rapidly and selectively induced by nerve growth factor in
RT   PC12 cells.";
RL   Mol. Cell. Biol. 11:2335-2349(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90005425; PubMed=2676516;
RA   Possenti R., Eldridge J.D., Paterson B.M., Grasso A., Levi A.;
RT   "A protein induced by NGF in PC12 cells is stored in secretory
RT   vesicles and released through the regulated pathway.";
RL   EMBO J. 8:2217-2223(1989).
RN   [3]
RP   REVISIONS.
RA   Possenti R.;
RL   Submitted (OCT-1995) to the SWISS-PROT data bank.
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92309005; PubMed=1377233;
RA   Hawley R.J., Scheibe R.J., Wagner J.A.;
RT   "NGF induces the expression of the VGF gene through a cAMP response
RT   element.";
RL   J. Neurosci. 12:2573-2581(1992).
CC   -!- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF CELL-CELL
CC       INTERACTIONS OR IN SYNATOGENESIS DURING THE MATURATION OF THE
CC       NERVOUS SYSTEM.
CC   -!- SUBCELLULAR LOCATION: SECRETORY VESICLES.
CC   -!- TISSUE SPECIFICITY: CENTRAL AND PERIPHERAL NERVOUS TISSUE.
CC   -!- INDUCTION: BY NGF.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M60522; AAA41700.1; -.
DR   EMBL; M60525; AAA86428.1; -.
DR   EMBL; M74223; AAA42336.1; -.
DR   PIR; S05381; S05381.
KW   Growth factor; Signal.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24    617       VGF PROTEIN.
FT   DOMAIN      439    452       ASP/GLU-RICH (ACIDIC).
FT   CONFLICT    342    342       A -> G (IN REF. 4).
SQ   SEQUENCE   617 AA;  68192 MW;  A41A447CD51F67B6 CRC64;


  Query Match             68.3%;  Score 41;  DB 1;  Length 617;
  Best Local Similarity   58.3%;  Pred. No. 3.1;
  Matches    7;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAE 12
           | |||:|::| |
Db     190 TRTHTLTRVNLE 201


RESULT   6
KNH2_BOVIN
ID   KNH2_BOVIN     STANDARD;      PRT;   619 AA.
AC   P01045;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Kininogen, HMW II precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84014106; PubMed=6571699;
RA   Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
RT   "A single gene for bovine high molecular weight and low molecular
RT   weight kininogens.";
RL   Nature 305:545-549(1983).
RN   [2]
RP   SEQUENCE OF 19-376.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
RN   [3]
RP   SEQUENCE OF 376-391.
RX   MEDLINE=70180420; PubMed=4986212;
RA   Kato H., Nagasawa S., Suzuki T.;
RT   "Studies on the structure of bovine kininogen: cleavages of disulfide
RT   bonds and of methionyl bonds in kininogen-II.";
RL   J. Biochem. 67:313-323(1970).
RN   [4]
RP   SEQUENCE OF 387-455.
RX   MEDLINE=76260155; PubMed=956151;
RA   Han Y.N., Kato H., Iwanaga S., Suzuki T.;
RT   "Primary structure of bovine plasma high-molecular-weight kininogen.
RT   The amino acid sequence of a glycopeptide portion (fragment 1)
RT   following the C-terminus ot the bradykinin moiety.";
RL   J. Biochem. 79:1201-1222(1976).
RN   [5]
RP   SEQUENCE OF 456-496.
RX   MEDLINE=75170265; PubMed=1169237;
RA   Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
RT   "Studies on the primary structure of bovine high-molecular-weight
RT   kininogen. Amino acid sequence of a fragment ('histidine-rich
RT   peptide') released by plasma kallikrein.";
RL   J. Biochem. 77:55-68(1975).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
```

```
CC       HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
CC       TO FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
CC       PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE
CC       PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
CC       A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
CC       MUSCLE CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C)
CC       NATRIURESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
CC       (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
CC       IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
CC       RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
CC       (4F) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
CC       ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
CC       ACTION).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS: HMW II AND LMW II KININOGEN PRECURSORS ARE
CC       PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
CC       SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
CC       TO RESIDUE 398.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V01492; CAA24736.1; -.
DR   EMBL; V01492; CAA24737.1; ALT_SEQ.
DR   PIR; A01282; KGBOH2.
DR   PIR; B29559; B29559.
DR   HSSP; P01038; 1A90.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Blood coagulation; Signal;
KW   Inflammatory response.
FT   SIGNAL       1    18
FT   CHAIN       19   619       KININOGEN, HMW II.
FT   CHAIN       19   376       HEAVY CHAIN.
FT   PEPTIDE    378   386       BRADYKININ.
FT   CHAIN      387   619       LIGHT CHAIN.
FT   DOMAIN      19   135       CYSTATIN-LIKE 1.
FT   DOMAIN     136   256       CYSTATIN-LIKE 2.
FT   DOMAIN     257   376       CYSTATIN-LIKE 3.
FT   MOD_RES     19    19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD    87    87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   136   136       O-LINKED (PARTIAL).
FT   CARBOHYD   168   168       N-LINKED (GLCNAC. . .); OR 169.
FT   CARBOHYD   197   197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD   204   204       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   280   280       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   400   400       O-LINKED.
FT   DISULFID    27   589       INTERCHAIN.
FT   DISULFID    82    93
FT   DISULFID   106   125
FT   DISULFID   141   144
FT   DISULFID   205   217
FT   DISULFID   228   247
FT   DISULFID   261   264
FT   DISULFID   325   337
FT   DISULFID   348   367
FT   VARIANT    398   398       T -> P.
FT   VARIANT    401   401       L -> V.
FT   VARIANT    454   454       H -> K.
SQ   SEQUENCE   619 AA;  68710 MW;  F04320A8EB0EE0DA CRC64;


  Query Match             68.3%;  Score 41;  DB 1;  Length 619;
  Best Local Similarity   72.7%;  Pred. No. 3.1;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAE 12
           | |:| |||||
Db     279 LNHSIAKLNAE 289


RESULT   7
MYSA_CAEEL
ID   MYSA_CAEEL     STANDARD;      PRT;  1969 AA.
AC   P12844;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Myosin heavy chain A (MHC A).
GN   MYO-3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=89178677; PubMed=2926820;
RA   Dibb N.J., Maruyama I.N., Krause M., Karn J.;
RT   "Sequence analysis of the complete Caenorhabditis elegans myosin
RT   heavy chain gene family.";
RL   J. Mol. Biol. 205:603-613(1989).
CC   -!- FUNCTION: MUSCLE CONTRACTION.
CC   -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC       HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC       AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC   -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC   -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC       CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC       CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC   -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC       ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC   -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC       MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
CC       SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC       SUBFRAGMENT (S2).
CC   -!- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN
CC       C.ELEGANS.
CC   -!- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY
CC       WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X08067; CAA30856.1; -.
DR   PIR; S02771; S02771.
DR   HSSP; P08799; 1LVK.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR004009; Myosin_N.
DR   InterPro; IPR002928; Myosin_tail.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF02736; Myosin_N; 1.
```

```
DR   Pfam; PF01576; Myosin_tail; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW   ATP-binding; Methylation; Alkylation; Multigene family.
FT   DOMAIN        1    793       MYOSIN HEAD-LIKE.
FT   DOMAIN      794    823       IQ.
FT   DOMAIN      857   1969       COILED COIL (POTENTIAL).
FT   NP_BIND     179    186       ATP.
FT   DOMAIN      667    689       ACTIN-BINDING.
FT   DOMAIN      770    784       ACTIN-BINDING.
FT   MOD_RES     130    130       METHYLATION (TRI-) (POTENTIAL).
FT   MOD_RES     707    707       ALKYLATION (SH-1) (POTENTIAL).
FT   MOD_RES     717    717       ALKYLATION (SH-2) (POTENTIAL).
SQ   SEQUENCE   1969 AA;  225509 MW;  64577BBAF7EAD80A CRC64;


  Query Match           61.7%;  Score 37;  DB 1;  Length 1969;
  Best Local Similarity   72.7%;  Pred. No. 63;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 LTHTITKLNAE 12
          ||  |:|||||
Db   1370 LTRQISKLNAE 1380


RESULT   8
SM3D_CHICK
ID   SM3D_CHICK     STANDARD;      PRT;   761 AA.
AC   Q90663;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Semaphorin 3D precursor (Collapsin-2) (COLL-2).
GN   SEMA3D OR COLL2.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95329269; PubMed=7605628;
RA   Luo Y., Shepherd I., Li J., Renzi M.J., Chang S., Raper J.A.;
RT   "A family of molecules related to collapsin in the embryonic chick
RT   nervous system.";
RL   Neuron 14:1131-1140(1995).
CC   -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
CC       CONES. COULD POTENTIALLY ACT AS REPULSIVE CUES TOWARD SPECIFIC
CC       NEURONAL POPULATIONS. BINDS TO NEUROPILIN.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: DEVELOPING SPINAL CORD AND DEVELOPING VISUAL
CC       SYSTEM. COLLAPSIN-1, -2, -3, AND -5 BIND TO OVERLAPPING BUT
CC       DISTINCT AXON TRACTS.
CC   -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
CC       THIRD OF THE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U28240; AAA86896.1; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003659; PSI.
DR   InterPro; IPR001627; Sema.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF01403; Sema; 1.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00423; PSI; 1.
KW   Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW   Developmental protein; Glycoprotein.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    761       SEMAPHORIN 3D.
FT   DOMAIN      245    543       SEMA.
FT   DOMAIN      646    726       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      727    757       ARG/LYS-RICH (BASIC).
FT   DISULFID    653    719       BY SIMILARITY.
FT   CARBOHYD    127    127       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    595    595       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   761 AA;  87300 MW;  3E09AE3DBA53F46B CRC64;


  Query Match           60.0%;  Score 36;  DB 1;  Length 761;
  Best Local Similarity   70.0%;  Pred. No. 35;
  Matches    7;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TLTHTITKLN 10
          |  ||| |||
Db    659 TFIHTIVKLN 668


RESULT   9
KF4A_MOUSE
ID   KF4A_MOUSE     STANDARD;      PRT;  1231 AA.
AC   P33174;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Chromosome-associated kinesin KIF4A (Chromokinesin).
GN   KIF4A OR KIF4 OR KNS4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE
RP   SPECIFICITY.
RC   STRAIN=ICR; TISSUE=Brain;
RX   MEDLINE=95014709; PubMed=7929562;
RA   Sekine Y., Okada Y., Noda Y., Kondo S., Aizawa H., Takemura R.,
RA   Hirokawa N.;
RT   "A novel microtubule-based motor protein (KIF4) for organelle
RT   transports, whose expression is regulated developmentally.";
RL   J. Cell Biol. 127:187-201(1994).
RN   [2]
RP   SEQUENCE OF 91-240 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=93077686; PubMed=1447303;
RA   Aizawa H., Sekine Y., Takemura R., Zhang Z., Nangaku M.,
RA   Hirokawa N.;
RT   "Kinesin family in murine central nervous system.";
RL   J. Cell Biol. 119:1287-1296(1992).
CC   -!- FUNCTION: REQUIRED FOR MITOTIC CHROMOSOMAL POSITIONING AND BIPOLAR
CC       SPINDLE STABILIZATION.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR, ASSOCIATED WITH MITOTIC
CC       CHROMOSOMES.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN PYRAMIDAL CELLS IN JUVENILE
CC       HIPPOCAMPUS, GRANULAR CELLS IN JUVENILE CEREBELLAR CORTEX AND IN
CC       ADULT SPLEEN.
CC   -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
CC       CHROMOKINESIN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D12646; BAA02167.1; -.
DR   PIR; D44259; D44259.
DR   HSSP; P17119; 3KAR.
DR   MGD; MGI:108389; Kif4.
DR   InterPro; IPR001752; kinesin.
DR   Pfam; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
DR   SMART; SM00129; KISc; 1.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW   Motor protein; Microtubules; ATP-binding; DNA-binding;
KW   Nuclear protein; Coiled coil.
FT   DOMAIN        1    350       KINESIN-MOTOR.
FT   DOMAIN      351   1000       COILED COIL (BY SIMILARITY).
FT   DOMAIN     1001   1231       GLOBULAR.
FT   NP_BIND      88     95       ATP (POTENTIAL).
FT   CONFLICT    112    112       I -> S (IN REF. 2).
SQ   SEQUENCE   1231 AA;  139551 MW;  F34F2C2D21158FE4 CRC64;


  Query Match             60.0%;  Score 36;  DB 1;  Length 1231;
  Best Local Similarity   66.7%;  Pred. No. 58;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 HTITKLNAE 12
           ||::||| |
Db     632 HTVSKLNQE 640


RESULT  10
YMEL_STRLN
ID   YMEL_STRLN     STANDARD;      PRT;   150 AA.
AC   P55049;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 17.2 kDa protein in MELC2-RNHH intergenic region (ORF3).
OS   Streptomyces lincolnensis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1915;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=78-11;
RA   Zhang H.Z., Piepersberg W.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X95703; CAA65001.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   150 AA;  17167 MW;  4F9B48B727D3967F CRC64;


  Query Match             58.3%;  Score 35;  DB 1;  Length 150;
  Best Local Similarity   60.0%;  Pred. No. 9.4;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLN 10
           |||| :| :|
Db      43 TLTHWVTSVN 52


RESULT  11
HSLV_XYLFA
ID   HSLV_XYLFA     STANDARD;      PRT;   178 AA.
AC   Q9PD95;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   ATP-dependent protease hslV (EC 3.4.25.-).
GN   HSLV OR XF1484.
OS   Xylella fastidiosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC   Xylella.
OX   NCBI_TaxID=2371;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=9A5C;
RX   MEDLINE=20365717; PubMed=10910347;
RA   Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA   Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA   Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA   Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA   Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA   Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA   Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA   Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA   Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA   Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA   Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA   Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA   Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA   Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA   Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA   Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA   Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA   de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA   Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA   Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA   de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA   da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA   da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA   de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
RA   Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA   Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT   "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL   Nature 406:151-159(2000).
CC   -!- FUNCTION: PROTEASE SUBUNIT OF A PROTEASOME-LIKE DEGRADATION
CC       COMPLEX (BY SIMILARITY).
CC   -!- SUBUNIT: INTERACTS WITH HSLU (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY T1B; ALSO KNOWN AS THE
CC       PROTEASOME B-TYPE FAMILY. HSLV SUBFAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AE003978; AAF84293.1; -.
DR   InterPro; IPR001353; Proteasome.
KW   Hydrolase; Protease; Complete proteome.
FT   ACT_SITE      8       8       BY SIMILARITY.
SQ   SEQUENCE   178 AA;  18803 MW;  A4406BED1A3275EE CRC64;


  Query Match             58.3%;  Score 35;  DB 1;  Length 178;
  Best Local Similarity   63.6%;  Pred. No. 11;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNA 11
```

```
          || ||: | ||
Db     28 TLGHTVMKSNA 38


RESULT 12
COAT_LSV
ID   COAT_LSV       STANDARD;      PRT;   291 AA.
AC   P27335;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   Coat protein (Capsid protein).
OS   Lily symptomless virus (LSV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Carlavirus.
OX   NCBI_TaxID=12173;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90218039; PubMed=2129538;
RA   Memelink J., van der Vlugt C.I.M., Linthorst H.J.M.,
RA   Derks A.F.L.M., Asjes C.J., Bol J.F.;
RT   "Homologies between the genomes of a carlavirus (lily symptomless
RT   virus) and a potexvirus (lily virus X) from lily plants.";
RL   J. Gen. Virol. 71:917-924(1990).
CC   -!- FUNCTION: SELF-ASSEMBLES WITH THE RNA TO FORM INFECTIOUS
CC       PARTICLES.
CC   -!- SIMILARITY: TO THE COAT PROTEINS FROM POTEXVIRUSES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15343; CAA33401.1; -.
DR   InterPro; IPR000052; Potex_carlavirus_coat.
DR   Pfam; PF00286; virus_P-coat; 1.
DR   PRINTS; PR00232; POTXCARLCOAT.
DR   ProDom; PD000603; Potex_carlavirus_coat; 1.
DR   PROSITE; PS00418; POTEX_CARLAVIRUS_COAT; 1.
KW   Coat protein.
SQ   SEQUENCE   291 AA;  32041 MW;  57E289F3EA726388 CRC64;


  Query Match             58.3%;  Score 35;  DB 1;  Length 291;
  Best Local Similarity   72.7%;  Pred. No. 19;
  Matches    8;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 LTHTITKLNAE 12
          ||  | |||||
Db     46 LTRLIEKLNAE 56


RESULT 13
KNG_RAT
ID   KNG_RAT        STANDARD;      PRT;   639 AA.
AC   P08934; P08933;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Kininogen precursor [Contains: Bradykinin].
GN   KNG.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RX   MEDLINE=87137443; PubMed=3029068;
RA   Kitagawa H., Kitamura N., Hayashida H., Miyata T., Nakanishi S.;
RT   "Differing expression patterns and evolution of the rat kininogen
RT   gene family.";
RL   J. Biol. Chem. 262:2190-2198(1987).
RN   [2]
RP   SEQUENCE FROM N.A. (LMW ISOFORM).
RX   MEDLINE=86008264; PubMed=2413018;
RA   Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT   "Primary structures of the mRNAs encoding the rat precursors for
RT   bradykinin and T-kinin. Structural relationship of kininogens with
RT   major acute phase protein and alpha 1-cysteine proteinase
RT   inhibitor.";
RL   J. Biol. Chem. 260:12054-12059(1985).
RN   [3]
RP   SEQUENCE OF 1-65 FROM N.A.
RC   STRAIN=BUFFALO;
RX   MEDLINE=87250580; PubMed=2439509;
RA   Fung W.-P., Schreiber G.;
RT   "Structure and expression of the genes for major acute phase alpha 1-
RT   protein (thiostatin) and kininogen in the rat.";
RL   J. Biol. Chem. 262:9298-9308(1987).
RN   [4]
RP   SEQUENCE OF 1-41 FROM N.A.
RC   STRAIN=WISTAR; TISSUE=Liver;
RX   MEDLINE=87137465; PubMed=3818598;
RA   Kageyama R., Kitamura N., Ohkubo H., Nakanishi S.;
RT   "Differing utilization of homologous transcription initiation sites
RT   of rat K and T kininogen genes under inflammation condition.";
RL   J. Biol. Chem. 262:2345-2351(1987).
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC       HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
CC       FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-
CC       INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
CC       BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
CC       PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
CC       CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
CC       DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
CC       MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
CC       PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
CC       OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
CC       A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
CC       INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
CC       KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
CC       CLOTTING.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- MISCELLANEOUS: RAT EXPRESS FOUR TYPES OF KININOGENS: THE CLASSICAL
CC       HMW/LMW KININOGENS AND TWO ADDITIONAL LMW-LIKE KININOGENS: T-I AND
CC       T-II.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29428; AAA41486.1; -.
DR   EMBL; M11884; AAA41487.1; -.
DR   EMBL; M14369; AAA41484.1; -.
DR   EMBL; M14369; AAA41485.1; ALT_SEQ.
DR   EMBL; M16455; AAA41482.1; -.
DR   PIR; A25486; A25486.
DR   PIR; A28055; A28055.
DR   HSSP; P01040; 1DVD.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
```

```
DR   PRINTS; PR00334; KININOGEN.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW   Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW   Alternative splicing; Multigene family.
FT   SIGNAL        1     18
FT   CHAIN        19    639       KININOGEN.
FT   CHAIN        19    380       KININOGEN HEAVY CHAIN.
FT   PEPTIDE     381    389       BRADYKININ.
FT   CHAIN       390    639       KININOGEN LIGHT CHAIN.
FT   DOMAIN       19    136       CYSTATIN-LIKE 1.
FT   DOMAIN      137    258       CYSTATIN-LIKE 2.
FT   DOMAIN      259    380       CYSTATIN-LIKE 3.
FT   DOMAIN      439    514       HIS-RICH.
FT   DISULFID     28    609       INTERCHAIN (BY SIMILARITY).
FT   DISULFID     83     94       BY SIMILARITY.
FT   DISULFID    107    126       BY SIMILARITY.
FT   DISULFID    142    145       BY SIMILARITY.
FT   DISULFID    206    218       BY SIMILARITY.
FT   DISULFID    229    248       BY SIMILARITY.
FT   DISULFID    264    267       BY SIMILARITY.
FT   DISULFID    328    340       BY SIMILARITY.
FT   DISULFID    351    370       BY SIMILARITY.
FT   CARBOHYD     82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    127    127       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    169    169       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    205    205       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    294    294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    529    529       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    402    433       VSPSYIARVQEERDPGNEQGPIHGHGWLHAKQ -> RLLNS
FT                                CEYKGRLLKAGAGPAPERQAEASTVTP (IN ISOFORM
FT                                LMW).
FT   VARSPLIC    434    639       MISSING (IN ISOFORM LMW).
FT   CONFLICT     61     61       E -> K (IN REF. 2).
SQ   SEQUENCE   639 AA;  70933 MW;  D3172DF94FF56AF5 CRC64;


  Query Match             58.3%;  Score 35;  DB 1;  Length 639;
  Best Local Similarity   63.6%;  Pred. No. 45;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 LTHTITKLNAE 12
          | |:| :||||
Db    282 LGHSIAQLNAE 292


RESULT  14
KNG_MOUSE
ID   KNG_MOUSE      STANDARD;      PRT;   661 AA.
AC   O08677; O08676;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Kininogen precursor [Contains: Bradykinin].
GN   KNG.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC   STRAIN=C57BL/6 x CBA; TISSUE=Liver;
RA   Takano M., Kondoh J., Yayama K., Okamoto H.;
RT   "Molecular cloning of cDNAs for mouse low- and high- molecular
RT   kininogen.";
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC       HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC       HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
CC       FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-
CC       INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
CC       BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
CC       PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
CC       CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
CC       DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
CC       MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
CC       PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
CC       OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
CC       A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
CC       INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
CC       LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
CC       KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
CC       CLOTTING (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC   -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D84435; BAA19743.1; -.
DR   EMBL; D84415; BAA19742.1; -.
DR   MGD; MGI:1097705; Kng.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 1.
KW   Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW   Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW   Alternative splicing.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    661       KININOGEN.
FT   CHAIN        19    379       KININOGEN HEAVY CHAIN.
FT   CHAIN       380    388       BRADYKININ.
FT   CHAIN       389    661       KININOGEN LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    257       CYSTATIN-LIKE 2.
FT   DOMAIN      258    379       CYSTATIN-LIKE 3.
FT   DOMAIN      439    524       HIS-RICH.
FT   DISULFID     28    631       INTERCHAIN (BY SIMILARITY).
FT   DISULFID     83     94       BY SIMILARITY.
FT   DISULFID    107    125       BY SIMILARITY.
FT   DISULFID    141    144       BY SIMILARITY.
FT   DISULFID    205    217       BY SIMILARITY.
FT   DISULFID    228    247       BY SIMILARITY.
FT   DISULFID    263    266       BY SIMILARITY.
FT   DISULFID    327    339       BY SIMILARITY.
FT   DISULFID    350    369       BY SIMILARITY.
FT   CARBOHYD     82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    242    242       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    401    432       VSPPYIAREQEERDAETEQGPTHGHGWLHEKQ -> RLLRA
FT                                CEYKGRLSKAGAEPAPERQAESSQVKQ (IN ISOFORM
FT                                LMW).
FT   VARSPLIC    433    661       MISSING (IN ISOFORM LMW).
SQ   SEQUENCE   661 AA;  73102 MW;  774460258D58796E CRC64;


  Query Match             58.3%;  Score 35;  DB 1;  Length 661;
  Best Local Similarity   63.6%;  Pred. No. 46;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy       2 LTHTITKLNAE 12
           | |:| :||||
Db     281 LGHSIAQLNAE 291


RESULT  15
ATH1_YEAST
ID   ATH1_YEAST     STANDARD;      PRT;  1211 AA.
AC   P48016;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Vacuolar acid trehalase precursor (EC 3.2.1.28) (Alpha,alpha-
DE   trehalase) (Alpha,alpha-trehalose glucohydrolase).
GN   ATH1 OR YPR026W OR YP9367.06.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE OF 130-1211 FROM N.A.
RC   STRAIN=MW303-1B;
RX   MEDLINE=96076626; PubMed=7502577;
RA   Destruelle M., Holzer H., Klionsky D.J.;
RT   "Isolation and characterization of a novel yeast gene, ATH1, that is
RT   required for vacuolar acid trehalase activity.";
RL   Yeast 11:1015-1025(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RA   Badcock K., Churcher C.M., Barrell B.G., Rajandream M.A.;
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Alpha,alpha-trehalose + H(2)O = 2 D-glucose.
CC   -!- SUBCELLULAR LOCATION: LYSOSOME-LIKE VACUOLES (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO FAMILY 65 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X84156; CAA58961.1; -.
DR   EMBL; Z49274; CAA89280.1; -.
DR   EMBL; Z71255; CAA95022.1; -.
DR   SGD; S0006230; ATH1.
KW   Hydrolase; Glycosidase; Glycoprotein; Signal.
FT   SIGNAL        1      ?       POTENTIAL.
FT   CHAIN         ?   1211       VACUOLAR ACID TREHALASE.
FT   CARBOHYD     28     28       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     32     32       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     98     98       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    207    207       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    238    238       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    247    247       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    255    255       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    259    259       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    325    325       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    370    370       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    376    376       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    488    488       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    539    539       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    568    568       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    628    628       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    638    638       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    696    696       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    705    705       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    879    879       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    897    897       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    910    910       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    972    972       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    990    990       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1031   1031       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1049   1049       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1064   1064       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1147   1147       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1157   1157       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT   1168   1211       LQVGDKGNTDWRKTRYIVVAVQGVYDDYDDDNKGATIKEIV
FT                                LND -> FAGG (IN REF. 1).
SQ   SEQUENCE   1211 AA;  136920 MW;  34AEC44B0B648DEC CRC64;


  Query Match             58.3%;  Score 35;  DB 1;  Length 1211;
  Best Local Similarity   70.0%;  Pred. No. 89;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAE 12
           ||:  |||||
Db     541 THSQAKLNAE 550


Search completed: July  1, 2002, 16:30:12
Job time: 621 sec
```

GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:29:44 ; Search time 75.26 Seconds
(without alignments)
27.584 Million cell updates/sec

Title: US-09-461-061A-2
Perfect score: 60
Sequence: 1 TLTHTITKLNAE 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 41 | 68.3 | 615 | 4 | Q9UDW8 | Q9udw8 homo sapien |
| 2 | 41 | 68.3 | 616 | 4 | O15240 | O15240 homo sapien |
| 3 | 39 | 65.0 | 265 | 2 | P95599 | P95599 rhizobium e |
| 4 | 39 | 65.0 | 534 | 16 | Q9PC17 | Q9pc17 xylella fas |
| 5 | 39 | 65.0 | 3010 | 12 | Q68833 | Q68833 hepatitis c |
| 6 | 39 | 65.0 | 3010 | 12 | Q9QIX4 | Q9qix4 hepatitis c |
| 7 | 39 | 65.0 | 3010 | 12 | Q9QIX3 | Q9qix3 hepatitis c |
| 8 | 38 | 63.3 | 268 | 3 | Q03984 | Q03984 saccharomyc |
| 9 | 38 | 63.3 | 498 | 3 | Q03983 | Q03983 saccharomyc |
| 10 | 38 | 63.3 | 3010 | 12 | Q9QIX8 | Q9qix8 hepatitis c |
| 11 | 38 | 63.3 | 3010 | 12 | Q9QIX7 | Q9qix7 hepatitis c |
| 12 | 37 | 61.7 | 103 | 2 | O84895 | O84895 salmonella |
| 13 | 37 | 61.7 | 786 | 5 | Q21673 | Q21673 caenorhabdi |
| 14 | 37 | 61.7 | 1992 | 5 | Q21440 | Q21440 caenorhabdi |
| 15 | 36 | 60.0 | 60 | 2 | Q9EVP2 | Q9evp2 escherichia |
| 16 | 36 | 60.0 | 65 | 9 | Q9MBZ6 | Q9mbz6 bacteriopha |
| 17 | 36 | 60.0 | 67 | 2 | Q9FCW6 | Q9fcw6 escherichia |
| 18 | 36 | 60.0 | 127 | 13 | Q90WA9 | Q90wa9 gallus gall |
| 19 | 36 | 60.0 | 197 | 16 | Q9K794 | Q9k794 bacillus ha |
| 20 | 36 | 60.0 | 305 | 16 | Q9PQ27 | Q9pq27 ureaplasma |
| 21 | 36 | 60.0 | 371 | 16 | Q929I2 | Q929i2 listeria in |
| 22 | 36 | 60.0 | 395 | 16 | Q55456 | Q55456 synechocyst |
| 23 | 36 | 60.0 | 399 | 2 | O31393 | O31393 bacillus me |
| 24 | 36 | 60.0 | 646 | 5 | Q9VPT4 | Q9vpt4 drosophila |
| 25 | 36 | 60.0 | 723 | 4 | O95789 | O95789 homo sapien |
| 26 | 36 | 60.0 | 1145 | 6 | Q9XSN8 | Q9xsn8 sus scrofa |
| 27 | 36 | 60.0 | 1267 | 10 | Q943D5 | Q943d5 oryza sativ |
| 28 | 36 | 60.0 | 2005 | 5 | Q9VFH5 | Q9vfh5 drosophila |
| 29 | 35.5 | 59.2 | 269 | 3 | Q00551 | Q00551 cryphonectr |
| 30 | 35 | 58.3 | 80 | 12 | Q81572 | Q81572 hepatitis c |
| 31 | 35 | 58.3 | 138 | 12 | Q68204 | Q68204 hepatitis c |
| 32 | 35 | 58.3 | 138 | 12 | Q68206 | Q68206 hepatitis c |
| 33 | 35 | 58.3 | 138 | 12 | Q68208 | Q68208 hepatitis c |
| 34 | 35 | 58.3 | 138 | 12 | Q68209 | Q68209 hepatitis c |
| 35 | 35 | 58.3 | 138 | 12 | Q68211 | Q68211 hepatitis c |
| 36 | 35 | 58.3 | 138 | 12 | Q68215 | Q68215 hepatitis c |
| 37 | 35 | 58.3 | 138 | 12 | Q68216 | Q68216 hepatitis c |
| 38 | 35 | 58.3 | 138 | 12 | Q68217 | Q68217 hepatitis c |
| 39 | 35 | 58.3 | 138 | 12 | Q68218 | Q68218 hepatitis c |
| 40 | 35 | 58.3 | 138 | 12 | Q68221 | Q68221 hepatitis c |
| 41 | 35 | 58.3 | 138 | 12 | Q68223 | Q68223 hepatitis c |
| 42 | 35 | 58.3 | 138 | 12 | Q68226 | Q68226 hepatitis c |
| 43 | 35 | 58.3 | 138 | 12 | Q68230 | Q68230 hepatitis c |
| 44 | 35 | 58.3 | 138 | 12 | Q68234 | Q68234 hepatitis c |
| 45 | 35 | 58.3 | 138 | 12 | Q68235 | Q68235 hepatitis c |

ALIGNMENTS

RESULT 1
Q9UDW8
ID Q9UDW8 PRELIMINARY; PRT; 615 AA.
AC Q9UDW8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE WUGSC:H_DJ0747G18.3 PROTEIN.
GN WUGSC:H_DJ0747G18.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99063792; PubMed=9847074;
RA Sulston J.E., Waterston R.;
RT "Toward a complete human genome sequence.";
RL Genome Res. 8:1097-1108(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Cordes M., Doela D.;
RT "The sequence of Homo sapiens PAC clone RP4-747G18.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AC004876; AAD45830.1; -.
SQ SEQUENCE 615 AA; 67257 MW; 198097C5622AC087 CRC64;

Query Match 68.3%; Score 41; DB 4; Length 615;
Best Local Similarity 58.3%; Pred. No. 14;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 TLTHTITKLNAE 12
          | |||:|::| |
Db    187 TRTHTLTRVNLE 198
```

```
RESULT   2
O15240
ID   O15240      PRELIMINARY;      PRT;   616 AA.
AC   O15240;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NEURO-ENDOCRINE SPECIFIC PROTEIN VGF.
GN   VGF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RX   MEDLINE=98008940; PubMed=9344675;
RA   Canu N., Possenti R., Ricco A.S., Rocchi M., Levi A.;
RT   "Cloning, structural organization analysis and chromosomal assignment
RT   of the human gene for neurosecretory protein VGF.";
RL   Genomics 45:443-446(1997).
DR   EMBL; Y12661; CAA73210.1; -.
SQ   SEQUENCE   616 AA;  67286 MW;  CD1920610201BEB9 CRC64;


  Query Match             68.3%;  Score 41;  DB 4;  Length 616;
  Best Local Similarity   58.3%;  Pred. No. 14;
  Matches    7;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAE 12
           | |||:|::| |
Db     187 TRTHTLTRVNLE 198


RESULT   3
P95599
ID   P95599      PRELIMINARY;      PRT;   265 AA.
AC   P95599;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   FIXK.
GN   FIXK.
OS   Rhizobium etli.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=29449;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CE-3;
RX   PubMed=10223993;
RA   Soberon M., Lopez O., Morera C., Girard M.L., Tabche M.L., Miranda J.;
RT   "Enhanced nitrogen fixation in a Rhizobium etli ntrC mutant that
RT   overproduces the Bradyrhizobium japonicum symbiotic terminal oxidase
RT   cbb3.";
RL   Appl. Environ. Microbiol. 65:2015-2019(1999).
CC   -!- SIMILARITY: BELONGS TO THE CRP/FNR FAMILY OF TRANSCRIPTIONAL
CC       REGULATORS.
DR   EMBL; U76906; AAC15887.1; -.
DR   InterPro; IPR000595; cNMP_binding.
DR   InterPro; IPR001808; HTH_CRP.
DR   Pfam; PF00027; cNMP_binding; 1.
DR   Pfam; PF00325; crp; 1.
DR   SMART; SM00100; cNMP; 1.
DR   SMART; SM00419; HTH_CRP; 1.
KW   DNA-binding; Transcription regulation.
SQ   SEQUENCE   265 AA;  29277 MW;  DA90B0819AEAA97B CRC64;


  Query Match             65.0%;  Score 39;  DB 2;  Length 265;
  Best Local Similarity   77.8%;  Pred. No. 14;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKL 9
           || ||:|||
Db     207 TLAHTVTKL 215


RESULT   4
Q9PC17
ID   Q9PC17      PRELIMINARY;      PRT;   534 AA.
AC   Q9PC17;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   METHYLTRANSFERASE.
GN   XF1968.
OS   Xylella fastidiosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC   Xylella.
OX   NCBI_TaxID=2371;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=9A5C;
RX   MEDLINE=20365717; PubMed=10910347;
RA   Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA   Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA   Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA   Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA   Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA   Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA   Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA   Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA   Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA   Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA   Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA   Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA   Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA   Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA   Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA   Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA   Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA   de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA   Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA   Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA   de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA   da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA   da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA   de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
RA   Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA   Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT   "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL   Nature 406:151-159(2000).
DR   EMBL; AE004016; AAF84770.1; -.
DR   InterPro; IPR002295; D21N6_mtfrase.
DR   InterPro; IPR001091; N4_MTase.
DR   InterPro; IPR002052; N6_Mtase.
DR   InterPro; IPR002941; N6_N4_Mtase.
DR   Pfam; PF01555; N6_N4_Mtase; 1.
DR   PRINTS; PR00506; D21N6MTFRASE.
DR   PRINTS; PR00508; S21N4MTFRASE.
DR   PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW   Complete proteome.
SQ   SEQUENCE   534 AA;  60730 MW;  3B429E738DB4E699 CRC64;


  Query Match             65.0%;  Score 39;  DB 16;  Length 534;
  Best Local Similarity   58.3%;  Pred. No. 30;
  Matches    7;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAE 12
           |  | : |||||
Db     359 TTAHAVAKLNAE 370
```

```
RESULT   5
Q68833
ID   Q68833      PRELIMINARY;      PRT;  3010 AA.
AC   Q68833;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)].
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HCV-L2;
RA   Cho J.M., Park Y.-W., Lee Y.-B., Yang J.-Y., Kim C.-H., Choo S.-H.,
RA   Ryu W.-S.;
RT   "Molecular cloning of hepatitis C virus genome from a single Korean
RT   blood donor.";
RL   Mol. Cells 5:317-324(1995).
CC   -!- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR   EMBL; U01214; AAA75355.1; -.
DR   HSSP; P26663; 1JXP.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR002518; HCV_NS2.
DR   InterPro; IPR004109; HCV_NS3.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; HCV_RdRP; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   ProDom; PD186062; HCV_NS1; 1.
KW   ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
KW   Nonstructural protein; Polyprotein; RNA-directed RNA polymerase;
KW   Transmembrane.
SQ   SEQUENCE   3010 AA;  326855 MW;  455BA4BF8CEF210E CRC64;


  Query Match             65.0%;  Score 39;  DB 12;  Length 3010;
  Best Local Similarity   88.9%;  Pred. No. 1.8e+02;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKL 9
           |||| ||||
Db    1636 TLTHPITKL 1644


RESULT   6
Q9QIX4
ID   Q9QIX4      PRELIMINARY;      PRT;  3010 AA.
AC   Q9QIX4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)].
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MD9-1;
RX   MEDLINE=20013325; PubMed=10544098;
RA   Nagayama K., Kurosaki M., Enomoto N., Maekawa S.y., Miyasaka Y.,
RA   Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT   "Time-related changes in full-length hepatitis C virus and hepatitis
RT   activity.";
RL   Virology 263:244-253(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MD9-1;
RA   Nagayama K., Kurosaki M., Enomoto N., Maekawa S., Miyasaka Y.,
RA   Sakamoto N., Fukuma T., Tazawa J., Izumi N., Marumo F., Sato C.;
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR   EMBL; AF165061; AAD56196.1; -.
DR   HSSP; P26663; 1NS3.
DR   MEROPS; S29.001; -.
DR   MEROPS; U39.001; -.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR002518; HCV_NS2.
DR   InterPro; IPR004109; HCV_NS3.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; HCV_RdRP; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   ProDom; PD186062; HCV_NS1; 1.
KW   ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
KW   Nonstructural protein; Polyprotein; RNA-directed RNA polymerase;
KW   Transmembrane.
SQ   SEQUENCE   3010 AA;  327386 MW;  94C94662C44A7695 CRC64;


  Query Match             65.0%;  Score 39;  DB 12;  Length 3010;
  Best Local Similarity   88.9%;  Pred. No. 1.8e+02;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKL 9
           |||| ||||
Db    1636 TLTHPITKL 1644


RESULT   7
Q9QIX3
ID   Q9QIX3      PRELIMINARY;      PRT;  3010 AA.
AC   Q9QIX3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)].
OS   Hepatitis C virus.
```

```
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MD9-2;
RX   MEDLINE=20013325; PubMed=10544098;
RA   Nagayama K., Kurosaki M., Enomoto N., Maekawa S.y., Miyasaka Y.,
RA   Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT   "Time-related changes in full-length hepatitis C virus and hepatitis
RT   activity.";
RL   Virology 263:244-253(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MD9-2;
RA   Nagayama K., Kurosaki M., Enomoto N., Maekawa S., Miyasaka Y.,
RA   Sakamoto N., Fukuma T., Tazawa J., Izumi N., Marumo F., Sato C.;
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR   EMBL; AF165062; AAD56197.1; -.
DR   HSSP; P26663; 1NS3.
DR   MEROPS; S29.001; -.
DR   MEROPS; U39.001; -.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR002518; HCV_NS2.
DR   InterPro; IPR004109; HCV_NS3.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; HCV_RdRP; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   ProDom; PD186062; HCV_NS1; 1.
KW   ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
KW   Nonstructural protein; Polyprotein; RNA-directed RNA polymerase;
KW   Transmembrane.
SQ   SEQUENCE   3010 AA;  327253 MW;  9F1B0B3F536774FA CRC64;


  Query Match          65.0%;  Score 39;  DB 12;  Length 3010;
  Best Local Similarity  88.9%;  Pred. No. 1.8e+02;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKL 9
            |||| ||||
Db    1636 TLTHPITKL 1644


RESULT   8
Q03984
ID   Q03984      PRELIMINARY;      PRT;   268 AA.
AC   Q03984;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL 31.1 KDA PROTEIN.
GN   YDR179W-A.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB972;
RA   Murphy L., Harris D.E.;
RL   Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB972;
RA   Barrell B., Rajandream M.A.;
RL   Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z46727; CAA86686.1; -.
DR   SGD; S0002587; YDR179W-A.
KW   Hypothetical protein.
SQ   SEQUENCE   268 AA;  31096 MW;  45C7F756F361C71D CRC64;


  Query Match          63.3%;  Score 38;  DB 3;  Length 268;
  Best Local Similarity  72.7%;  Pred. No. 23;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNA 11
            || | |:||||
Db     144 TLQHWISKLNA 154


RESULT   9
Q03983
ID   Q03983      PRELIMINARY;      PRT;   498 AA.
AC   Q03983;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL 57.9 KDA PROTEIN.
GN   YDR179W-A.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB972;
RA   Murphy L., Harris D.E.;
RL   Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB972;
RA   Barrell B., Rajandream M.A.;
RL   Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z46727; CAA86685.1; -.
DR   SGD; S0002587; YDR179W-A.
KW   Hypothetical protein.
SQ   SEQUENCE   498 AA;  57916 MW;  60959427D7A230CD CRC64;


  Query Match          63.3%;  Score 38;  DB 3;  Length 498;
  Best Local Similarity  72.7%;  Pred. No. 43;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNA 11
            || | |:||||
Db     374 TLQHWISKLNA 384


RESULT  10
Q9QIX8
ID   Q9QIX8      PRELIMINARY;      PRT;  3010 AA.
AC   Q9QIX8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
```

```
DE   (NS1)].
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MD7-1;
RX   MEDLINE=20013325; PubMed=10544098;
RA   Nagayama K., Kurosaki M., Enomoto N., Maekawa S.y., Miyasaka Y.,
RA   Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT   "Time-related changes in full-length hepatitis C virus and hepatitis
RT   activity.";
RL   Virology 263:244-253(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MD7-1;
RA   Nagayama K., Kurosaki M., Enomoto N., Maekawa S., Miyasaka Y.,
RA   Sakamoto N., Fukuma T., Tazawa J., Izumi N., Marumo F., Sato C.;
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR   EMBL; AF165057; AAD56192.1; -.
DR   HSSP; P26663; 1JXP.
DR   MEROPS; S29.001; -.
DR   MEROPS; U39.001; -.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR002518; HCV_NS2.
DR   InterPro; IPR004109; HCV_NS3.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; HCV_RdRP; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   ProDom; PD186062; HCV_NS1; 1.
KW   ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
KW   Nonstructural protein; Polyprotein; RNA-directed RNA polymerase;
KW   Transmembrane.
SQ   SEQUENCE   3010 AA;  326986 MW;  55F505A208C6E5CD CRC64;


  Query Match             63.3%;  Score 38;  DB 12;  Length 3010;
  Best Local Similarity   77.8%;  Pred. No. 2.7e+02;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKL 9
           |||| :|||
Db    1636 TLTHPVTKL 1644


RESULT  11
Q9QIX7
ID   Q9QIX7      PRELIMINARY;      PRT;  3010 AA.
AC   Q9QIX7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)].
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MD7-2;
RX   MEDLINE=20013325; PubMed=10544098;
RA   Nagayama K., Kurosaki M., Enomoto N., Maekawa S.y., Miyasaka Y.,
RA   Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT   "Time-related changes in full-length hepatitis C virus and hepatitis
RT   activity.";
RL   Virology 263:244-253(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MD7-2;
RA   Nagayama K., Kurosaki M., Enomoto N., Maekawa S., Miyasaka Y.,
RA   Sakamoto N., Fukuma T., Tazawa J., Izumi N., Marumo F., Sato C.;
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR   EMBL; AF165058; AAD56193.1; -.
DR   HSSP; P26663; 1JXP.
DR   MEROPS; S29.001; -.
DR   MEROPS; U39.001; -.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR002518; HCV_NS2.
DR   InterPro; IPR004109; HCV_NS3.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; HCV_RdRP; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   ProDom; PD186062; HCV_NS1; 1.
KW   ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
KW   Nonstructural protein; Polyprotein; RNA-directed RNA polymerase;
KW   Transmembrane.
SQ   SEQUENCE   3010 AA;  326974 MW;  A3556D74F0C3AD2B CRC64;


  Query Match             63.3%;  Score 38;  DB 12;  Length 3010;
  Best Local Similarity   77.8%;  Pred. No. 2.7e+02;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKL 9
           |||| :|||
Db    1636 TLTHPVTKL 1644


RESULT  12
O84895
ID   O84895      PRELIMINARY;      PRT;   103 AA.
AC   O84895;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   HYPOTHETICAL 11.3 KDA PROTEIN.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
```

```
OC   Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2;
RA   Figueroa-Bossi N., Bossi L.;
RT   "Prophage genes in Salmonella.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF001386; AAC26074.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   103 AA;  11325 MW;  84379F22D7906527 CRC64;


  Query Match             61.7%;  Score 37;  DB 2;  Length 103;
  Best Local Similarity   58.3%;  Pred. No. 13;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TLTHTITKLNAE 12
          || ||: | :||
Db     44 TLAHTVEKRDAE 55


RESULT  13
Q21673
ID   Q21673      PRELIMINARY;      PRT;   786 AA.
AC   Q21673;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   R03G8.4 PROTEIN.
GN   R03G8.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Coles L.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z69794; CAA93681.1; -.
DR   MEROPS; M01.UPW; -.
DR   InterPro; IPR001930; Aladiptase.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01433; Peptidase_M1; 1.
DR   PRINTS; PR00756; ALADIPTASE.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ   SEQUENCE   786 AA;  90448 MW;  CAA1ACD3A1803C9A CRC64;


  Query Match             61.7%;  Score 37;  DB 5;  Length 786;
  Best Local Similarity   63.6%;  Pred. No. 1.1e+02;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 LTHTITKLNAE 12
          ||| | ||| :
Db    636 LTHAIAKLNQD 646


RESULT  14
Q21440
ID   Q21440      PRELIMINARY;      PRT;  1992 AA.
AC   Q21440;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update).
DE   K12F2.1 PROTEIN.
GN   K12F2.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Harris B.R.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z78199; CAB01576.1; -.
DR   HSSP; P08799; 1MND.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR001609; myosin_head.
DR   InterPro; IPR004009; Myosin_N.
DR   InterPro; IPR002928; Myosin_tail.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 2.
DR   Pfam; PF02736; Myosin_N; 1.
DR   Pfam; PF01576; Myosin_tail; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
SQ   SEQUENCE   1992 AA;  228362 MW;  8676AA34DAD7E835 CRC64;


  Query Match             61.7%;  Score 37;  DB 5;  Length 1992;
  Best Local Similarity   72.7%;  Pred. No. 2.8e+02;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 LTHTITKLNAE 12
          ||  |:|||||
Db   1393 LTRQISKLNAE 1403


RESULT  15
Q9EVP2
ID   Q9EVP2      PRELIMINARY;      PRT;    60 AA.
AC   Q9EVP2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL 7.2 KDA PROTEIN.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H.I.8.;
RX   MEDLINE=20407286; PubMed=10948097;
RA   Unkmeir A., Schmidt H.;
RT   "Structural analysis of phage-borne stx genes and their flanking
RT   sequences in shiga toxin-producing escherichia coli and shigella
RT   dysenteriae type 1 strains.";
RL   Infect. Immun. 68:4856-4864(2000).
DR   EMBL; AJ271139; CAC05572.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   60 AA;  7199 MW;  37845BB5F08D1EFE CRC64;


  Query Match             60.0%;  Score 36;  DB 2;  Length 60;
  Best Local Similarity   70.0%;  Pred. No. 12;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TLTHTITKLN 10
```

||| |:|| |
Db 46 TLTKTLTKFN 55

Search completed: July 1, 2002, 16:29:45
Job time: 699 sec

GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:19:43 ; Search time 95.97 Seconds
(without alignments)
13.889 Million cell updates/sec

Title: US-09-461-061A-2
Perfect score: 60
Sequence: 1 TLTHTITKLNAE 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 60 | 100.0 | 12 | 21 | AAY95406 | Anti-angiogenic pe |
| 2 | 60 | 100.0 | 32 | 21 | AAY95408 | Anti-angiogenic D3 |
| 3 | 60 | 100.0 | 117 | 14 | AAR33350 | Domaine 3, bradyki |
| 4 | 60 | 100.0 | 122 | 21 | AAB37447 | Human kininogen D3 |
| 5 | 60 | 100.0 | 123 | 21 | AAY95426 | Human high mol.wt. |
| 6 | 60 | 100.0 | 248 | 22 | ABG21102 | Novel human diagno |
| 7 | 60 | 100.0 | 369 | 22 | ABG21099 | Novel human diagno |
| 8 | 60 | 100.0 | 435 | 22 | ABG21105 | Novel human diagno |
| 9 | 60 | 100.0 | 644 | 22 | ABG21101 | Novel human diagno |
| 10 | 42 | 70.0 | 436 | 5 | AAP40257 | Bradykinin protein |
| 11 | 41 | 68.3 | 434 | 5 | AAP40633 | Bradykinin protein |
| 12 | 41 | 68.3 | 616 | 22 | AAU09069 | Human neuroendocri |
| 13 | 41 | 68.3 | 669 | 22 | AAO12939 | Human polypeptide |
| 14 | 38 | 63.3 | 16 | 21 | AAY95409 | Anti-angiogenic D3 |
| 15 | 38 | 63.3 | 210 | 21 | AAB54127 | Human pancreatic c |
| 16 | 37 | 61.7 | 70 | 22 | AAM06387 | Human foetal prote |
| 17 | 37 | 61.7 | 70 | 22 | AAM06833 | Human foetal prote |
| 18 | 37 | 61.7 | 121 | 21 | AAB33188 | Pinus radiata tran |
| 19 | 37 | 61.7 | 327 | 22 | AAG91497 | C glutamicum prote |
| 20 | 36 | 60.0 | 80 | 22 | AAU53927 | Propionibacterium |
| 21 | 36 | 60.0 | 646 | 22 | ABB60240 | Drosophila melanog |
| 22 | 36 | 60.0 | 723 | 21 | AAB41943 | Human ORFX ORF1707 |
| 23 | 36 | 60.0 | 750 | 22 | ABG09135 | Novel human diagno |
| 24 | 36 | 60.0 | 750 | 22 | ABG10240 | Novel human diagno |
| 25 | 36 | 60.0 | 750 | 22 | ABG14540 | Novel human diagno |
| 26 | 36 | 60.0 | 750 | 22 | ABG19900 | Novel human diagno |
| 27 | 36 | 60.0 | 1145 | 18 | AAW32097 | Miniature swine re |
| 28 | 36 | 60.0 | 1145 | 22 | AAB73286 | Retoviral protein |
| 29 | 36 | 60.0 | 1316 | 22 | AAB27248 | Maize RAD50. Zea |
| 30 | 36 | 60.0 | 2005 | 22 | ABB59854 | Drosophila melanog |
| 31 | 35 | 58.3 | 63 | 13 | AAR25862 | HCV polypeptide 9. |
| 32 | 35 | 58.3 | 63 | 14 | AAW41740 | Hepatitis C virus |
| 33 | 35 | 58.3 | 87 | 13 | AAR25884 | HK9. Hepatitis C |
| 34 | 35 | 58.3 | 111 | 22 | AAO00443 | Human polypeptide |
| 35 | 35 | 58.3 | 113 | 22 | AAO00391 | Human polypeptide |
| 36 | 35 | 58.3 | 194 | 12 | AAR11277 | Hepatitis C virus |
| 37 | 35 | 58.3 | 194 | 15 | AAR52737 | HCV antigen. Synt |
| 38 | 35 | 58.3 | 194 | 16 | AAR66618 | Hepatitis C virus |
| 39 | 35 | 58.3 | 194 | 16 | AAR66625 | Hepatitis C virus |
| 40 | 35 | 58.3 | 194 | 16 | AAR66626 | Hepatitis C virus |
| 41 | 35 | 58.3 | 194 | 20 | AAY14973 | HCV J1 NS3 domain |
| 42 | 35 | 58.3 | 194 | 20 | AAY01617 | Proein encoded by |
| 43 | 35 | 58.3 | 194 | 20 | AAY01612 | Protein encoded by |
| 44 | 35 | 58.3 | 194 | 20 | AAW30586 | Hepatitis C virus |
| 45 | 35 | 58.3 | 195 | 13 | AAR25863 | HCV polypeptide 10 |

ALIGNMENTS

RESULT 1
AAY95406
ID AAY95406 standard; Peptide; 12 AA.
XX
AC AAY95406;
XX
DT 25-SEP-2000 (first entry)
XX
DE Anti-angiogenic peptide N-terminal fragment.
XX
KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW therapy; human; D3 peptide.
XX
OS Homo sapiens.
XX
PN WO200035407-A2.
XX
PD 22-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28465.
XX
PR 16-DEC-1998; 98US-0112427.
XX
PA (UTEM ) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI McCrae RK;
XX
DR WPI; 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,

bad date

PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT 3 analog -
XX
PS Claim 3; Page 25; 44pp; English.
XX
CC The present sequence is that of an N-terminal fragment of a novel
CC anti-angiogenic D3 peptide (see AAY95408) derived from human high
CC mol.wt. kininogen (HK) domain 3 (see AAY95426). The full-length D3
CC peptide inhibits endothelial cell proliferation and thus possesses
CC anti-angiogenic activity. It is an example of peptides of the
CC invention (see AAY95405-26) that are analogues of certain sites in
CC the HK domain 3. The peptides inhibit endothelial cell proliferation
CC and may also induce endothelial cell apoptosis. Compositions
CC including the peptides are used in claimed methods for inhibiting
CC angiogenesis, inhibiting endothelial cell proliferation, and
CC inducing endothelial cell apoptosis. Cancer, rheumatoid arthritis,
CC and ocular disorders characterized by undesired vascularization of
CC the retina are treated.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 60; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNAE 12
||||||||||||
Db 1 tlthtitklnae 12

RESULT 2
AAY95408
ID AAY95408 standard; Peptide; 32 AA.
XX
AC AAY95408;
XX
DT 25-SEP-2000 (first entry)
XX
DE Anti-angiogenic D3 peptide.
XX
KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW therapy; human; D3 peptide.
XX
OS Homo sapiens.
XX
PN WO200035407-A2.
XX
PD 22-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28465.
XX
PR 16-DEC-1998; 98US-0112427.
XX
PA (UTEM ) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI McCrae RK;
XX
DR WPI; 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,
PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT 3 analog -
XX
PS Claim 4; Page 26; 44pp; English.
XX
CC The present sequence is that of a D3 peptide derived from human
CC high mol.wt. kininogen (HK) domain 3 (see AAY95426). The D3 peptide
CC inhibits endothelial cell proliferation and thus possesses
CC anti-angiogenic activity. It is an example of D3 peptides of the
CC invention (see AAY95405-26) that are analogues of certain sites in
CC the HK domain 3, in this case amino acid residues Asn275-Lys282.
CC The peptides inhibit endothelial cell proliferation and may also
CC induce endothelial cell apoptosis. Compositions including the
CC peptides are used in claimed methods for inhibiting angiogenesis,
CC inhibiting endothelial cell proliferation, and inducing endothelial
CC cell apoptosis. Cancer, rheumatoid arthritis, and ocular disorders
CC characterized by undesired vascularization of the retina are treated.
XX
SQ Sequence 32 AA;

Query Match 100.0%; Score 60; DB 21; Length 32;
Best Local Similarity 100.0%; Pred. No. 0.00097;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNAE 12
||||||||||||
Db 1 tlthtitklnae 12

RESULT 3
AAR33350
ID AAR33350 standard; protein; 117 AA.
XX
AC AAR33350;
XX
DT 01-JUL-1993 (first entry)
XX
DE Domaine 3, bradykinin release activating peptide.
XX
KW Domain 3; human; kininogen; heavy chain; low molecular weight; plasma;
KW trypsin; platelet; activation; granule contents; hemostasis; thrombin;
KW tissue plasminogen activator;thrombosis; inflammatory response;
KW endothelial cell; von Willebrand factor;
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..18
FT /note= "Leader peptide"
FT Protein 19..117
FT /note= "Mature protein"
XX
PN WO9303748-A.
XX
PD 04-MAR-1993.
XX
PF 13-AUG-1992; 92WO-US06809.
XX
PR 13-AUG-1991; 91US-0744545.
XX
PA (UTEM ) UNIV TEMPLE.
XX
PI Jiang Y, Schmaier AB;
XX
DR WPI; 1993-093714/11.
XX
PT Use of trypsin-cleavage fragment of human kininogen - for
PT increasing vascular bradykinin release, for lowering blood
PT pressure and treating hypertension
XX
PS Disclosure; Fig 1; 46pp; English.
XX
CC The sequence given represents domain 3, amino acids 246-362, of
CC the human kininogen heavy chain. Domain 3 was isolated from low
CC molecular weight kininogen, derived from human plasma, by cleavage
CC with trypsin. Domain 3 peptide inhibits platelet activation causing
CC a marked decrease in the platelets ability to aggregate and secrete
CC their granule contents. The granule contents comprise proteins which
CC participate in hemostasis, thrombosis and the inflammatory response.
CC Domain 3 also inhibits endothelial cell activation shown by a decrease

```
CC   in secretion of endothelial cell contents such as tissue plasminogen
CC   activator and von Willebrand factor.  Domain 3 functions to inhibit
CC   cell activation by blocking thrombin binding to its target cells, the
CC   peptide is a selective inhibitor of thrombin-induced platlet
CC   activation.
XX
SQ   Sequence   117 AA;


  Query Match             100.0%;  Score 60;  DB 14;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 0.0044;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAE 12
           ||||||||||||
Db      18 tlthtitklnae 29


RESULT   4
AAB37447
ID   AAB37447 standard; protein; 122 AA.
XX
AC   AAB37447;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human kininogen D3.
XX
KW   Enzyme; legumain; endopeptidase; cystatin; human; kininogen.
XX
OS   Homo sapiens.
XX
PN   WO200064945-A1.
XX
PD   02-NOV-2000.
XX
PF   20-APR-2000; 2000WO-GB01571.
XX
PR   22-APR-1999;   99GB-0009133.
XX
PA   (BABR-) BABRAHAM INST.
XX
PI   Abrahamson M,  Barrett AJ;
XX
DR   WPI; 2000-687316/67.
XX
PT   Inhibition of mammalian legumain or legumain-related endopeptidase by
PT   cystatin involves interaction with second papain-non-reactive site of
PT   cystatin  -
XX
PS   Disclosure; Fig 4; 45pp; English.
XX
CC   The present invention relates to inhibition of the enzymatic activity of
CC   legumain or a legumain-related endopeptidase by cystatin. The inhibition
CC   involves an interaction between legumain and a papain-non-reactive site
CC   of cystatin. Legumain (EC 3.4.22.34) is a cysteine endopeptidase, and
CC   performs a protein-processing function. The present sequence is human
CC   kininogen D3, which was used in the present invention. Kininogen is a
CC   type 3 cystatin.
XX
SQ   Sequence   122 AA;


  Query Match             100.0%;  Score 60;  DB 21;  Length 122;
  Best Local Similarity   100.0%;  Pred. No. 0.0046;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAE 12
           ||||||||||||
Db      23 tlthtitklnae 34


RESULT   5
AAY95426
ID   AAY95426 standard; Peptide; 123 AA.
XX
AC   AAY95426;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Human high mol.wt. kininogen domain 3.
XX
KW   Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW   endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW   rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW   therapy; human; D3 peptide.
XX
OS   Homo sapiens.
XX
PN   WO200035407-A2.
XX
PD   22-JUN-2000.
XX
PF   02-DEC-1999;   99WO-US28465.
XX
PR   16-DEC-1998;   98US-0112427.
XX
PA   (UTEM ) UNIV TEMPLE.
PA   (MCCR/) MCCRAE R K.
XX
PI   McCrae RK;
XX
DR   WPI; 2000-442247/38.
XX
PT   Composition for inhibiting angiogenesis and endothelial cell
PT   proliferation, inducing endothelial cell apoptosis and treating cancer,
PT   rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT   3 analog  -
XX
PS   Disclosure; Page 4; 44pp; English.
XX
CC   The present sequence is that of domain 3 of human high mol.wt.
CC   kininogen (HK).  The invention provides peptides (see AAY95405-24)
CC   that are analogues of certain sites in the HK domain 3,
CC   specifically Asn275-Lys282, Cys246-Cys249, Leu331-Tyr338 and
CC   Tyr299-Ser314.  The peptides, in which native Cys residues may be
CC   replaced by Ala residues, inhibit endothelial cell proliferation
CC   and may also induce endothelial cell apoptosis.  Compositions
CC   including the peptides are used in claimed methods for inhibiting
CC   angiogenesis, inhibiting endothelial cell proliferation, and
CC   inducing endothelial cell apoptosis.  Cancer, rheumatoid arthritis,
CC   and ocular disorders characterized by undesired vascularization of
CC   the retina are treated.
XX
SQ   Sequence   123 AA;


  Query Match             100.0%;  Score 60;  DB 21;  Length 123;
  Best Local Similarity   100.0%;  Pred. No. 0.0047;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAE 12
           ||||||||||||
Db      29 tlthtitklnae 40


RESULT   6
ABG21102
ID   ABG21102 standard; Protein; 248 AA.
XX
AC   ABG21102;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #21093.
XX
```

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS85289.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 51461; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 248 AA;

Query Match 100.0%; Score 60; DB 22; Length 248;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNAE 12
||||||||||||
Db 78 tlthtitklnae 89

RESULT 7
ABG21099
ID ABG21099 standard; Protein; 369 AA.
XX
AC ABG21099;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #21090.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS85286.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 51458; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 369 AA;

Query Match 100.0%; Score 60; DB 22; Length 369;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNAE 12
||||||||||||
Db 321 tlthtitklnae 332

RESULT 8
ABG21105
ID ABG21105 standard; Protein; 435 AA.
XX
AC ABG21105;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #21096.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

```
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS85292.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 51464; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   435 AA;


  Query Match             100.0%;  Score 60;  DB 22;  Length 435;
  Best Local Similarity   100.0%;  Pred. No. 0.02;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TLTHTITKLNAE 12
          ||||||||||||
Db    285 tlthtitklnae 296


RESULT   9
ABG21101
ID   ABG21101 standard; Protein; 644 AA.
XX
AC   ABG21101;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #21092.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS85288.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 51460; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   644 AA;


  Query Match             100.0%;  Score 60;  DB 22;  Length 644;
  Best Local Similarity   100.0%;  Pred. No. 0.032;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TLTHTITKLNAE 12
          ||||||||||||
Db    281 tlthtitklnae 292


RESULT  10
AAP40257
ID   AAP40257 standard; Protein; 436 AA.
XX
AC   AAP40257;
XX
DT   30-JUL-1992  (first entry)
XX
DE   Bradykinin protein precursor: type I (pKG13, pKG59).
XX
KW   Blood pressure; kininogen; probe.
XX
FH   Key             Location/Qualifiers
FT   Peptide         380..388
```

```
FT                   /label= bradykinin
FT   Peptide         393..397
FT                   /note= "probe (AAN40241)-encoded sequence"
XX
PN   JP59125896-A.
XX
PD   20-JUL-1984.
XX
PF   07-JAN-1983;   83JP-0000984.
XX
PR   07-JAN-1983;   83JP-0000984.
XX
PA   (MITU ) MITSUBISHI CHEM IND KK.
XX
DR   WPI; 1984-216122/35.
DR   N-PSDB; AAN40242.
XX
PT   c-Dna fragment of protein precursor - used to code bradykinin
XX
PS   Disclosure; Fig 2; 6 pp; Japanese.
XX
CC   Bradykinin is a peptide consisting of nine amino acids.  It has the
CC   biological effect of decreasing blood pressure.  Although kininogen
CC   is known as a protein-precursor of bradykinin, its structure is unknown
CC   because of the difficulty in collecting large enough samples of
CC   kininogen for structural investigation.
XX
SQ   Sequence   436 AA;


  Query Match             70.0%;  Score 42;  DB 5;  Length 436;
  Best Local Similarity   72.7%;  Pred. No. 29;
  Matches    8;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAE 12
           |:|:| |||||
Db     281 lshsiaklnae 291


RESULT  11
AAP40633
ID   AAP40633 standard; Protein; 434 AA.
XX
AC   AAP40633;
XX
DT   30-JUL-1992  (first entry)
XX
DE   Bradykinin protein precursor: type II (pKG146, pKG254).
XX
KW   Blood pressure; kininogen; probe..
XX
FH   Key             Location/Qualifiers
FT   Peptide         378..386
FT                   /label= bradykinin
FT   Peptide         391..395
FT                   /note= "probe (AAN40241)-encoded sequence"
XX
PN   JP59125896-A.
XX
PD   20-JUL-1984.
XX
PF   07-JAN-1983;   83JP-0000984.
XX
PR   07-JAN-1983;   83JP-0000984.
XX
PA   (MITU ) MITSUBISHI CHEM IND KK.
XX
DR   WPI; 1984-216122/35.
DR   N-PSDB; AAN40314.
XX
PT   c-Dna fragment of protein precursor - used to code bradykinin
XX
PS   Disclosure; Fig 2; 6 pp; Japanese.
XX
CC   Bradykinin is a peptide consisting of nine amino acids.  It has the
CC   biological effect of decreasing blood pressure.  Although kininogen
CC   is known as a protein-precursor of bradykinin, its structure is unknown
CC   because of the difficulty in collecting large enough samples of
CC   kininogen for structural investigation.
XX
SQ   Sequence   434 AA;


  Query Match             68.3%;  Score 41;  DB 5;  Length 434;
  Best Local Similarity   72.7%;  Pred. No. 43;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAE 12
           | |:| |||||
Db     279 lnhsiaklnae 289


RESULT  12
AAU09069
ID   AAU09069 standard; Protein; 616 AA.
XX
AC   AAU09069;
XX
DT   19-DEC-2001  (first entry)
XX
DE   Human neuroendocrine VGF.
XX
KW   Human; long-term memory protein; LTM; neuroendocrine VGF;
KW   neuroleptic; anticonvulsant; nootropic; neuroprotective; C/EBPbeta;
KW   cerebroprotective; drug discovery; therapeutic profiling;
KW   learning disability; memory impairment; brain injury; epilepsy;
KW   mental retardation; senile dementia; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
PN   WO200174298-A2.
XX
PD   11-OCT-2001.
XX
PF   02-APR-2001; 2001WO-US10661.
XX
PR   31-MAR-2000; 2000US-193614P.
XX
PA   (UYBR-) UNIV BROWN RESEACH FOUND.
PA   (HUGH-) HUGHES HOWARD MED INST.
XX
PI   Alberini CM,  Bear MF;
XX
DR   WPI; 2001-626335/72.
DR   N-PSDB; AAS14697.
XX
PT   Regulating memory consolidation in an animal comprising treating with
PT   an agent that modulates activity of one or more genes from zif268,
PT   insulin-like growth factor, glutamate receptor 2, c/EBPbeta and VGF -
XX
PS   Disclosure; Page 98-100; 100pp; English.
XX
CC   The invention relates to modulating long term memory consolidation in an
CC   animal comprises treating with an agent that modulates the activity of
CC   one or more of genes from zif268, insulin-like growth factor (IGF),
CC   glutamate receptor 1 (GluR1), glutamate receptor 2 (GluR2), c/EBPbeta
CC   and neuroendocrine VGF (neurotropin-inducible gene). The method is useful
CC   for identifying an agent which modulates memory consolidation. The method
CC   is useful for conducting a drug and/or target discovery business, which
CC   comprises conducting therapeutic profiling of the agents (or their
CC   analogues) identified, for efficacy and toxicity in animals, and
CC   formulating a pharmaceutical preparation including one or more agents
CC   identified as having an acceptable therapeutic profile and/or licensing
CC   to a third party the rights for further drug development of the
CC   identified agents. The method of conducting drug discovery business
CC   further comprises an additional step of establishing a distribution
```

CC system for distributing the preparation for sale and may optionally
CC include establishing a sales group for marketing the preparation. A
CC pharmaceutical composition containing the agent is useful for enhancing
CC memory consolidation in an animal, or for augmenting learning and memory,
CC or otherwise for enhancing the functional performance of central nervous
CC system neurons, where the agent is a cAMP elevating agent (agonist)
CC preferably a cAMP analogue or cAMP phosphodiesterase inhibitor, which
CC activates adenylate cyclase. The composition is useful for treating
CC diseases associated with learning disabilities, memory impairment e.g.
CC due to toxicant exposure, brain injury, epilepsy, mental retardation in
CC children and senile dementia, including Alzheimer's disease. The
CC present sequence represents human neuroendocrine VGF.
XX
SQ Sequence 616 AA;

Query Match 68.3%; Score 41; DB 22; Length 616;
Best Local Similarity 58.3%; Pred. No. 64;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNAE 12
| |||:|::| |
Db 187 trthtltrvnle 198

RESULT 13
AAO12939
ID AAO12939 standard; Protein; 669 AA.
XX
AC AAO12939;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 26831.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR N-PSDB; AAI92870.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
PS Claim 20; SEQ ID NO 26831; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 669 AA;

Query Match 68.3%; Score 41; DB 22; Length 669;
Best Local Similarity 58.3%; Pred. No. 71;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNAE 12
| |||:|::| |
Db 239 trthtltrvnle 250

RESULT 14
AAY95409
ID AAY95409 standard; Peptide; 16 AA.
XX
AC AAY95409;
XX
DT 25-SEP-2000 (first entry)
XX
DE Anti-angiogenic D3 peptide.
XX
KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW therapy; human; D3 peptide.
XX
OS Homo sapiens.
XX
PN WO200035407-A2.
XX
PD 22-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28465.
XX
PR 16-DEC-1998; 98US-0112427.
XX
PA (UTEM ) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI McCrae RK;
XX
DR WPI; 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,
PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT 3 analog -
XX
PS Claim 6; Page 26; 44pp; English.
XX
CC The present sequence is that of a D3 peptide derived from human
CC high mol.wt. kininogen (HK) domain 3 (see AAY95426). The D3 peptide
CC inhibits endothelial cell proliferation and thus possesses
CC anti-angiogenic activity. It is an example of D3 peptides of the
CC invention (see AAY95405-26) that are analogues of certain sites in
CC the HK domain 3, in this case amino acid residues Asn275-Lys282.
CC The peptides inhibit endothelial cell proliferation and may also
CC induce endothelial cell apoptosis. Compositions including the
CC peptides are used in claimed methods for inhibiting angiogenesis,
CC inhibiting endothelial cell proliferation, and inducing endothelial
CC cell apoptosis. Cancer, rheumatoid arthritis, and ocular disorders
CC characterized by undesired vascularization of the retina are treated.
CC The IC50 value for the present peptide was less than 0.8 uM for
CC inhibition of fibroblast growth factor-induced HUVEC cell
CC proliferation.
XX

```
SQ   Sequence   16 AA;


  Query Match             63.3%;  Score 38;  DB 21;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 3.1;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 TITKLNAE 12
          ||||||||
Db      1 titklnae 8


RESULT  15
AAB54127
ID   AAB54127 standard; Protein; 210 AA.
XX
AC   AAB54127;
XX
DT   09-MAR-2001  (first entry)
XX
DE   Human pancreatic cancer antigen protein sequence SEQ ID NO:579.
XX
KW   Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
KW   detection; diagnosis; identification; cytostatic; neuroprotective;
KW   nootropic; immunomodulatory; relaxant; contraceptive; gynaecological;
KW   antiinflammatory; cardiant; gene therapy; chromosome mapping;
KW   linkage analysis; tissue identification; tissue typing; forensic;
KW   neural; immune system; muscular; reproductive; gastrointestinal;
KW   pulmonary; cardiovascular; renal; proliferative.
XX
OS   Homo sapiens.
XX
PN   WO200055320-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05989.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-579444/54.
DR   N-PSDB; AAC98892.
XX
PT   New nucleic acid that is a pancreatic cancer antigen for preventing,
PT   treating, or ameliorating a medical condition, particular pancreatic
PT   cancer, or for use in assays for diagnosing a pathological condition -
XX
PS   Claim 11; Page 1015-1016; 1379pp; English.
XX
CC   AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC   proteins, called pancreatic cancer antigens, given in AAB54008 to
CC   AAB54466. The human pancreatic cancer antigens have cytostatic,
CC   neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
CC   gynaecological, cardiant and antiinflammatory activities, and can be used
CC   in gene therapy. The polynucleotide and proteins can be used for
CC   preventing, treating, or ameliorating a medical condition or in assays
CC   for diagnosing a pathological condition or a susceptibility to one in a
CC   subject. Binding partners to the proteins and the activity of the
CC   proteins can be identified. The pancreatic cancer antigens can be used to
CC   detect, treat or prevent pancreatic disorders, especially cancer.
CC   Agonists and antagonists to the antigens can be screened for. The
CC   pancreatic cancer antigen polynucleotides can be used to design nucleic
CC   acid hybridisation probes that can be used in chromosome mapping, linkage
CC   analysis, tissue identification and/or typing and a variety of forensic
CC   and diagnostic methods. The proteins can be used to generate antibodies
CC   which are used to purify, detect and target the polypeptides, including
CC   both in vivo and in vitro diagnostic and therapeutic methods. The
CC   proteins can be used to treat or prevent neural, immune system, muscular,
CC   reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC   proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC   sequences used in the exemplification of the present invention.
XX
SQ   Sequence   210 AA;


  Query Match             63.3%;  Score 38;  DB 21;  Length 210;
  Best Local Similarity   58.3%;  Pred. No. 61;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TLTHTITKLNAE 12
          |||| :  :|||
Db    183 tlthyvrpinae 194


Search completed: July  1, 2002, 16:19:43
Job time: 147 sec
```

GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:18:00 ; Search time 35.15 Seconds
(without alignments)
8.339 Million cell updates/sec

Title: US-09-461-061A-2
Perfect score: 60
Sequence: 1 TLTHTITKLNAE 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 60 | 100.0 | 117 | 1 | US-08-193-114B-1 | Sequence 1, Appli |
| 2 | 60 | 100.0 | 117 | 5 | PCT-US92-06809-1 | Sequence 1, Appli |
| 3 | 35 | 58.3 | 189 | 2 | US-08-483-695-46 | Sequence 46, Appl |
| 4 | 35 | 58.3 | 189 | 2 | US-07-965-285-46 | Sequence 46, Appl |
| 5 | 35 | 58.3 | 189 | 2 | US-08-487-231-46 | Sequence 46, Appl |
| 6 | 35 | 58.3 | 189 | 4 | US-09-201-912-46 | Sequence 46, Appl |
| 7 | 35 | 58.3 | 247 | 1 | US-08-324-977-44 | Sequence 44, Appl |
| 8 | 35 | 58.3 | 247 | 2 | US-08-384-616-44 | Sequence 44, Appl |
| 9 | 35 | 58.3 | 247 | 2 | US-08-904-686A-44 | Sequence 44, Appl |
| 10 | 35 | 58.3 | 247 | 4 | US-09-315-850-44 | Sequence 44, Appl |
| 11 | 35 | 58.3 | 313 | 2 | US-08-483-695-7 | Sequence 7, Appli |
| 12 | 35 | 58.3 | 313 | 2 | US-07-965-285-7 | Sequence 7, Appli |
| 13 | 35 | 58.3 | 313 | 2 | US-08-487-231-7 | Sequence 7, Appli |
| 14 | 35 | 58.3 | 313 | 4 | US-09-201-912-7 | Sequence 7, Appli |
| 15 | 35 | 58.3 | 585 | 2 | US-08-867-941-21 | Sequence 21, Appl |
| 16 | 35 | 58.3 | 585 | 4 | US-09-074-658-21 | Sequence 21, Appl |
| 17 | 35 | 58.3 | 631 | 1 | US-08-700-356-1 | Sequence 1, Appli |
| 18 | 35 | 58.3 | 631 | 2 | US-08-936-865-1 | Sequence 1, Appli |
| 19 | 35 | 58.3 | 631 | 4 | US-09-128-314-2 | Sequence 2, Appli |
| 20 | 35 | 58.3 | 632 | 4 | US-09-198-723A-23 | Sequence 23, Appl |
| 21 | 35 | 58.3 | 638 | 4 | US-09-288-391-25 | Sequence 25, Appl |
| 22 | 35 | 58.3 | 646 | 4 | US-09-198-723A-60 | Sequence 60, Appl |
| 23 | 35 | 58.3 | 646 | 4 | US-09-198-723A-63 | Sequence 63, Appl |
| 24 | 35 | 58.3 | 646 | 4 | US-09-198-723A-66 | Sequence 66, Appl |
| 25 | 35 | 58.3 | 646 | 4 | US-09-198-723A-69 | Sequence 69, Appl |
| 26 | 35 | 58.3 | 646 | 4 | US-09-198-723A-72 | Sequence 72, Appl |
| 27 | 35 | 58.3 | 666 | 4 | US-09-198-723A-11 | Sequence 11, Appl |
| 28 | 35 | 58.3 | 666 | 4 | US-09-198-723A-12 | Sequence 12, Appl |
| 29 | 35 | 58.3 | 666 | 4 | US-09-198-723A-13 | Sequence 13, Appl |
| 30 | 35 | 58.3 | 666 | 4 | US-09-198-723A-14 | Sequence 14, Appl |
| 31 | 35 | 58.3 | 666 | 4 | US-09-198-723A-15 | Sequence 15, Appl |
| 32 | 35 | 58.3 | 666 | 4 | US-09-198-723A-16 | Sequence 16, Appl |
| 33 | 35 | 58.3 | 666 | 4 | US-09-198-723A-17 | Sequence 17, Appl |
| 34 | 35 | 58.3 | 666 | 4 | US-09-198-723A-18 | Sequence 18, Appl |
| 35 | 35 | 58.3 | 672 | 4 | US-09-198-723A-19 | Sequence 19, Appl |
| 36 | 35 | 58.3 | 672 | 4 | US-09-198-723A-20 | Sequence 20, Appl |
| 37 | 35 | 58.3 | 753 | 2 | US-08-867-941-20 | Sequence 20, Appl |
| 38 | 35 | 58.3 | 753 | 4 | US-09-074-658-20 | Sequence 20, Appl |
| 39 | 35 | 58.3 | 985 | 2 | US-08-867-941-13 | Sequence 13, Appl |
| 40 | 35 | 58.3 | 985 | 2 | US-08-867-941-17 | Sequence 17, Appl |
| 41 | 35 | 58.3 | 985 | 4 | US-09-074-658-13 | Sequence 13, Appl |
| 42 | 35 | 58.3 | 985 | 4 | US-09-074-658-17 | Sequence 17, Appl |
| 43 | 35 | 58.3 | 1000 | 2 | US-08-867-941-12 | Sequence 12, Appl |
| 44 | 35 | 58.3 | 1000 | 2 | US-08-867-941-16 | Sequence 16, Appl |
| 45 | 35 | 58.3 | 1000 | 4 | US-09-074-658-12 | Sequence 12, Appl |

ALIGNMENTS

```
RESULT 1
US-08-193-114B-1
; Sequence 1, Application US/08193114B
; Patent No. 5472945
;  GENERAL INFORMATION:
;    APPLICANT:  Schmaier, Alvin H.
;    APPLICANT:  Jiang, Yongping
;    TITLE OF INVENTION:  Modulation of Blood
;    TITLE OF INVENTION:  Pressure and Inhibition of Platelet Activation
;    TITLE OF INVENTION:  with Kininogen Fragment
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seidel, Gonda, Lavorgna &
;      ADDRESSEE:  Monaco, P.C.
;      STREET:  1800 Two Penn Center Plaza
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/193,114B
;      FILING DATE:  9 February 1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  U.S. Application
;      APPLICATION NUMBER:  Serial No. 5472945 07/744,545
;      FILING DATE:  13 August 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-137 CI1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  No. 5472945e
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  117 amino acids
;      TYPE:  peptide
;      TOPOLOGY:  linear
US-08-193-114B-1


  Query Match           100.0%;  Score 60;  DB 1;  Length 117;
  Best Local Similarity 100.0%;  Pred. No. 0.00095;
```

```
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TLTHTITKLNAE 12
          ||||||||||||
Db     18 TLTHTITKLNAE 29


RESULT   2
PCT-US92-06809-1
; Sequence 1, Application PC/TUS9206809
;  GENERAL INFORMATION:
;    APPLICANT:  Schmaier, Alvin H.
;    APPLICANT:  Jiang, Yongping
;    TITLE OF INVENTION:  Modulation of Blood
;    TITLE OF INVENTION:  Pressure by Altering Bradykinin Levels
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of the
;      ADDRESSEE:  Commonwealth System of Higher Education
;      STREET:  406 University Services
;      STREET:  Building
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/06809
;      FILING DATE:  19910813
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  U.S. Application
;      APPLICATION NUMBER:  Serial No.  744,545
;      FILING DATE:  13 August 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-137
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  117 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
PCT-US92-06809-1


  Query Match            100.0%;  Score 60;  DB 5;  Length 117;
  Best Local Similarity  100.0%;  Pred. No. 0.00095;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TLTHTITKLNAE 12
          ||||||||||||
Db     18 TLTHTITKLNAE 29


RESULT   3
US-08-483-695-46
; Sequence 46, Application US/08483695
; Patent No. 5866139
;  GENERAL INFORMATION:
;    APPLICANT:  Brechot, Christian
;    APPLICANT:  Kremsdorf, Dina
;    APPLICANT:  Porchon, Colette
;    TITLE OF INVENTION:  Nucleotide and Peptide Sequences of a
;    TITLE OF INVENTION:  Hepatitis C Virus Isolate, Diagnostic and Therapeutic
;    TITLE OF INVENTION:  Applications
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,695
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/965,285
;      FILING DATE:  18-MAR-1993
;      APPLICATION NUMBER:  FR 91 06 882
;      FILING DATE:  06-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meyers, Kenneth J.
;      REGISTRATION NUMBER:  25,146
;      REFERENCE/DOCKET NUMBER:  05286-0001-00000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-408-4000
;      TELEFAX:  202-408-4400
;  INFORMATION FOR SEQ ID NO:  46:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  189 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-483-695-46


  Query Match            58.3%;  Score 35;  DB 2;  Length 189;
  Best Local Similarity  87.5%;  Pred. No. 42;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TLTHTITK 8
          |||| |||
Db    182 TLTHPITK 189


RESULT   4
US-07-965-285-46
; Sequence 46, Application US/07965285
; Patent No. 5879904
;  GENERAL INFORMATION:
;    APPLICANT:  Brechot, Christian
;    APPLICANT:  Kremsdorf, Dina
;    APPLICANT:  Porchon, Colette
;    TITLE OF INVENTION:  Nucleotide and Peptide Sequences of a
;    TITLE OF INVENTION:  Hepatitis C Virus Isolate, Diagnostic and Therapeutic
;    TITLE OF INVENTION:  Applications
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/965,285
;      FILING DATE:  18-MAR-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 91 06 882
;      FILING DATE:  06-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meyers, Kenneth J.
;      REGISTRATION NUMBER:  25,146
;      REFERENCE/DOCKET NUMBER:  05286-0001-00000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-408-4000
;      TELEFAX:  202-408-4400
;  INFORMATION FOR SEQ ID NO:  46:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  189 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-07-965-285-46


  Query Match             58.3%;  Score 35;  DB 2;  Length 189;
  Best Local Similarity   87.5%;  Pred. No. 42;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTITK 8
           |||| |||
Db     182 TLTHPITK 189


RESULT   5
US-08-487-231-46
; Sequence 46, Application US/08487231
; Patent No. 5919454
;  GENERAL INFORMATION:
;    APPLICANT:  Brechot, Christian
;    APPLICANT:  Kremsdorf, Dina
;    APPLICANT:  Porchon, Colette
;    TITLE OF INVENTION:  Nucleotide and Peptide Sequences of a
;    TITLE OF INVENTION:  Hepatitis C Virus Isolate, Diagnostic and Therapeutic
;    TITLE OF INVENTION:  Applications
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,231
;      FILING DATE:  07-JUNE-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/965,285
;      FILING DATE:  18-MAR-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 91 06 882
;      FILING DATE:  06-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meyers, Kenneth J.
;      REGISTRATION NUMBER:  25,146
;      REFERENCE/DOCKET NUMBER:  05286-0001-02000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-408-4000
;      TELEFAX:  202-408-4400
;  INFORMATION FOR SEQ ID NO:  46:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  189 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-487-231-46


  Query Match             58.3%;  Score 35;  DB 2;  Length 189;
  Best Local Similarity   87.5%;  Pred. No. 42;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTITK 8
           |||| |||
Db     182 TLTHPITK 189


RESULT   6
US-09-201-912-46
; Sequence 46, Application US/09201912
; Patent No. 6210962
;  GENERAL INFORMATION:
;    APPLICANT:  Brechot, Christian
;    APPLICANT:  Kremsdorf, Dina
;    APPLICANT:  Porchon, Colette
;    TITLE OF INVENTION:  Nucleotide and Peptide Sequences of a
;    TITLE OF INVENTION:  Hepatitis C Virus Isolate, Diagnostic and Therapeutic
;    TITLE OF INVENTION:  Applications
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/201,912
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/965,285
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meyers, Kenneth J.
;      REGISTRATION NUMBER:  25,146
;      REFERENCE/DOCKET NUMBER:  05286-0001-00000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-408-4000
;      TELEFAX:  202-408-4400
;  INFORMATION FOR SEQ ID NO:  46:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  189 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-201-912-46


  Query Match             58.3%;  Score 35;  DB 4;  Length 189;
  Best Local Similarity   87.5%;  Pred. No. 42;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy       1 TLTHTITK 8
            |||| |||
Db     182 TLTHPITK 189


RESULT   7
US-08-324-977-44
; Sequence 44, Application US/08324977
; Patent No. 5747339
;  GENERAL INFORMATION:
;    APPLICANT:  OKAYAMA, Hiroto
;    APPLICANT:  FUKE, Isao
;    APPLICANT:  MORI, Chisato
;    APPLICANT:  TAKAMIZAWA, Akahisa
;    APPLICANT:  YOSHIDA, Iwao
;    TITLE OF INVENTION:  NON-A, NON-B HEPATITIS VIRUS GENOMIC
;    TITLE OF INVENTION:  cDNA AND ANTIGEN POLYPEPTIDE
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Armstrong, Westerman, Hattori, McLeland &
;      ADDRESSEE:  Naughton
;      STREET:  1725 K St. N.W. Suite 1000
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20006
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 in, 1.44Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS, Version 5.0
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/324,977
;      FILING DATE:  18-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 2-167466
;      FILING DATE:  25-JUN-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 2-230921
;      FILING DATE:  31-AUG-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 2-305605
;      FILING DATE:  09-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/099,706
;      FILING DATE:  30-JUL-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/769,996
;      FILING DATE:  02-OCT-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/635,451
;      FILING DATE:  28-DEC-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stevens-Smith, Theresa M.
;      REGISTRATION NUMBER:  36,281
;      REFERENCE/DOCKET NUMBER:  900703D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 659-2930
;      TELEFAX:  (202) 887-0357
;      TELEX:  440142
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  247 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-324-977-44


  Query Match           58.3%;  Score 35;  DB 1;  Length 247;
  Best Local Similarity  87.5%;  Pred. No. 57;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTITK 8
            |||| |||
Db      21 TLTHPITK 28


RESULT   8
US-08-384-616-44
; Sequence 44, Application US/08384616
; Patent No. 5847101
;  GENERAL INFORMATION:
;    APPLICANT:  OKAYAMA, Hiroto
;    APPLICANT:  FUKE, Isao
;    APPLICANT:  MORI, Chisato
;    APPLICANT:  TAKAMIZAWA, Akahisa
;    APPLICANT:  YOSHIDA, Iwao
;    TITLE OF INVENTION:  NON-A, NON-B HEPATITIS VIRUS GENOMIC
;    TITLE OF INVENTION:  cDNA AND ANTIGEN POLYPEPTIDE
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Armstrong, Westerman, Hattori, McLeland &
;      ADDRESSEE:  Naughton
;      STREET:  1725 K St. N.W. Suite 1000
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20006
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 in, 1.44Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS, Version 5.0
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/384,616
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/769,996
;      FILING DATE:  02-OCT-1991
;      APPLICATION NUMBER:  JP 2-167466
;      FILING DATE:  25-JUN-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 2-230921
;      FILING DATE:  31-AUG-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 2-305605
;      FILING DATE:  09-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/635,451
;      FILING DATE:  28-DEC-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stevens-Smith, Theresa M.
;      REGISTRATION NUMBER:  36,281
;      REFERENCE/DOCKET NUMBER:  900703B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 659-2930
;      TELEFAX:  (202) 887-0357
;      TELEX:  440142
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  247 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-384-616-44


  Query Match           58.3%;  Score 35;  DB 2;  Length 247;
  Best Local Similarity  87.5%;  Pred. No. 57;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTITK 8
            |||| |||
Db      21 TLTHPITK 28
```

```
RESULT   9
US-08-904-686A-44
; Sequence 44, Application US/08904686A
; Patent No. 5998130
;  GENERAL INFORMATION:
;    APPLICANT:  OKAYAMA, Hiroto
;    APPLICANT:  FUKE, Isao
;    APPLICANT:  MORI, Chisato
;    APPLICANT:  TAKAMIZAWA, Akahisa
;    APPLICANT:  YOSHIDA, Iwao
;    TITLE OF INVENTION:  NON-A, NON-B HEPATITIS VIRUS GENOMIC
;    TITLE OF INVENTION:  cDNA AND ANTIGEN POLYPEPTIDE
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Armstrong, Westerman, Hattori, McLeland &
;      ADDRESSEE:  Naughton
;      STREET:  1725 K St. N.W. Suite 1000
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20006
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 in, 1.44Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS, Version 5.0
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/904,686A
;      FILING DATE:  01-AUG-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/324,977
;      FILING DATE:  18-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 2-167466
;      FILING DATE:  25-JUN-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 2-230921
;      FILING DATE:  31-AUG-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 2-305605
;      FILING DATE:  09-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/099,706
;      FILING DATE:  30-JUL-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/769,996
;      FILING DATE:  02-OCT-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/635,451
;      FILING DATE:  28-DEC-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McLeland, Le-Nhung
;      REGISTRATION NUMBER:  31,541
;      REFERENCE/DOCKET NUMBER:  900703G
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 659-2930
;      TELEFAX:  (202) 887-0357
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  247 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-904-686A-44


  Query Match           58.3%;  Score 35;  DB 2;  Length 247;
  Best Local Similarity  87.5%;  Pred. No. 57;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTITK 8
                    ||||  |||
Db      21 TLTHPITK 28


RESULT  10
US-09-315-850-44
; Sequence 44, Application US/09315850
; Patent No. 6217872
;  GENERAL INFORMATION:
;    APPLICANT:  OKAYAMA, Hiroto
;    APPLICANT:  FUKE, Isao
;    APPLICANT:  MORI, Chisato
;    APPLICANT:  TAKAMIZAWA, Akahisa
;    APPLICANT:  YOSHIDA, Iwao
;    TITLE OF INVENTION:  NON-A, NON-B HEPATITIS VIRUS GENOMIC
;    TITLE OF INVENTION:  cDNA AND ANTIGEN POLYPEPTIDE
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Armstrong, Westerman, Hattori, McLeland &
;      ADDRESSEE:  Naughton
;      STREET:  1725 K St. N.W. Suite 1000
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20006
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 in, 1.44Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS, Version 5.0
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/315,850
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/904,686
;      FILING DATE:  01-AUG-1997
;      APPLICATION NUMBER:  US 08/324,977
;      FILING DATE:  18-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 2-167466
;      FILING DATE:  25-JUN-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 2-230921
;      FILING DATE:  31-AUG-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 2-305605
;      FILING DATE:  09-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/099,706
;      FILING DATE:  30-JUL-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/769,996
;      FILING DATE:  02-OCT-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/635,451
;      FILING DATE:  28-DEC-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McLeland, Le-Nhung
;      REGISTRATION NUMBER:  31,541
;      REFERENCE/DOCKET NUMBER:  900703G
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 659-2930
;      TELEFAX:  (202) 887-0357
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  247 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-315-850-44


  Query Match           58.3%;  Score 35;  DB 4;  Length 247;
```

```
  Best Local Similarity   87.5%;  Pred. No. 57;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTITK 8
           |||| |||
Db      21 TLTHPITK 28


RESULT  11
US-08-483-695-7
; Sequence 7, Application US/08483695
; Patent No. 5866139
;  GENERAL INFORMATION:
;    APPLICANT:  Brechot, Christian
;    APPLICANT:  Kremsdorf, Dina
;    APPLICANT:  Porchon, Colette
;    TITLE OF INVENTION:  Nucleotide and Peptide Sequences of a
;    TITLE OF INVENTION:  Hepatitis C Virus Isolate, Diagnostic and Therapeutic
;    TITLE OF INVENTION:  Applications
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,695
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/965,285
;      FILING DATE:  18-MAR-1993
;      APPLICATION NUMBER:  FR 91 06 882
;      FILING DATE:  06-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meyers, Kenneth J.
;      REGISTRATION NUMBER:  25,146
;      REFERENCE/DOCKET NUMBER:  05286-0001-00000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-408-4000
;      TELEFAX:  202-408-4400
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  313 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-483-695-7


  Query Match             58.3%;  Score 35;  DB 2;  Length 313;
  Best Local Similarity   87.5%;  Pred. No. 73;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTITK 8
           |||| |||
Db     182 TLTHPITK 189


RESULT  12
US-07-965-285-7
; Sequence 7, Application US/07965285
; Patent No. 5879904
;  GENERAL INFORMATION:
;    APPLICANT:  Brechot, Christian
;    APPLICANT:  Kremsdorf, Dina
;    APPLICANT:  Porchon, Colette
;    TITLE OF INVENTION:  Nucleotide and Peptide Sequences of a
;    TITLE OF INVENTION:  Hepatitis C Virus Isolate, Diagnostic and Therapeutic
;    TITLE OF INVENTION:  Applications
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/965,285
;      FILING DATE:  18-MAR-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 91 06 882
;      FILING DATE:  06-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meyers, Kenneth J.
;      REGISTRATION NUMBER:  25,146
;      REFERENCE/DOCKET NUMBER:  05286-0001-00000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-408-4000
;      TELEFAX:  202-408-4400
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  313 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-07-965-285-7


  Query Match             58.3%;  Score 35;  DB 2;  Length 313;
  Best Local Similarity   87.5%;  Pred. No. 73;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTITK 8
           |||| |||
Db     182 TLTHPITK 189


RESULT  13
US-08-487-231-7
; Sequence 7, Application US/08487231
; Patent No. 5919454
;  GENERAL INFORMATION:
;    APPLICANT:  Brechot, Christian
;    APPLICANT:  Kremsdorf, Dina
;    APPLICANT:  Porchon, Colette
;    TITLE OF INVENTION:  Nucleotide and Peptide Sequences of a
;    TITLE OF INVENTION:  Hepatitis C Virus Isolate, Diagnostic and Therapeutic
;    TITLE OF INVENTION:  Applications
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/487,231
;       FILING DATE:  07-JUNE-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/965,285
;       FILING DATE:  18-MAR-1993
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  FR 91 06 882
;       FILING DATE:  06-JUN-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Meyers, Kenneth J.
;       REGISTRATION NUMBER:  25,146
;       REFERENCE/DOCKET NUMBER:  05286-0001-02000
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-408-4000
;       TELEFAX:  202-408-4400
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  313 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-487-231-7


  Query Match             58.3%;  Score 35;  DB 2;  Length 313;
  Best Local Similarity   87.5%;  Pred. No. 73;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTITK 8
           |||| |||
Db     182 TLTHPITK 189


RESULT  14
US-09-201-912-7
; Sequence 7, Application US/09201912
; Patent No. 6210962
;  GENERAL INFORMATION:
;     APPLICANT:  Brechot, Christian
;     APPLICANT:  Kremsdorf, Dina
;     APPLICANT:  Porchon, Colette
;     TITLE OF INVENTION:  Nucleotide and Peptide Sequences of a
;     TITLE OF INVENTION:  Hepatitis C Virus Isolate, Diagnostic and Therapeutic
;     TITLE OF INVENTION:  Applications
;     NUMBER OF SEQUENCES:  46
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;       ADDRESSEE:  Dunner
;       STREET:  1300 I Street, N.W.
;       CITY:  Washington
;       STATE:  DC
;       COUNTRY:  USA
;       ZIP:  20005-3315
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/201,912
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/965,285
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Meyers, Kenneth J.
;       REGISTRATION NUMBER:  25,146
;       REFERENCE/DOCKET NUMBER:  05286-0001-00000
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-408-4000
;       TELEFAX:  202-408-4400
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  313 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-09-201-912-7


  Query Match             58.3%;  Score 35;  DB 4;  Length 313;
  Best Local Similarity   87.5%;  Pred. No. 73;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTITK 8
           |||| |||
Db     182 TLTHPITK 189


RESULT  15
US-08-867-941-21
; Sequence 21, Application US/08867941
; Patent No. 5977337
;  GENERAL INFORMATION:
;     APPLICANT:  Loosmore, Sheena M
;     APPLICANT:  Du, Run-Pan
;     APPLICANT:  Wang, Quijun
;     APPLICANT:  Yang, Yan-Ping
;     APPLICANT:  Klein, Michel H
;     TITLE OF INVENTION:  LACTOFERRIN RECEPTOR GENES OF MORAXELLA
;     NUMBER OF SEQUENCES:  67
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Sim & McBurney
;       STREET:  6th Floor, 330 University Avenue
;       CITY:  Toronto
;       STATE:  Ontario
;       COUNTRY:  Canada
;       ZIP:  M5G 1R7
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/867,941
;       FILING DATE:  03-JUN-1997
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Stewart, Michael I
;       REGISTRATION NUMBER:  24,973
;       REFERENCE/DOCKET NUMBER:  1038-681 MIS:jb
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (416) 595-1155
;       TELEFAX:  (416) 595-1163
;   INFORMATION FOR SEQ ID NO:  21:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  585 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-867-941-21


  Query Match             58.3%;  Score 35;  DB 2;  Length 585;
  Best Local Similarity   87.5%;  Pred. No. 1.5e+02;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTITK 8
           |||||| |
```

Db 510 TLTHTIPK 517

Search completed: July 1, 2002, 16:18:00
Job time: 44 sec

GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:17:16 ; Search time 35.15 Seconds
(without alignments)
5.559 Million cell updates/sec

Title: US-09-461-061A-1
Perfect score: 45
Sequence: 1 NNATFYFK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 45 | 100.0 | 117 | 1 | US-08-193-114B-1 | Sequence 1, Appli |
| 2 | 45 | 100.0 | 117 | 5 | PCT-US92-06809-1 | Sequence 1, Appli |
| 3 | 39 | 86.7 | 26 | 4 | US-08-676-242-15 | Sequence 15, Appl |
| 4 | 31 | 68.9 | 193 | 1 | US-08-483-140-30 | Sequence 30, Appl |
| 5 | 31 | 68.9 | 193 | 2 | US-08-485-938A-36 | Sequence 36, Appl |
| 6 | 31 | 68.9 | 441 | 1 | US-08-470-187-8 | Sequence 8, Appli |
| 7 | 31 | 68.9 | 441 | 1 | US-08-318-905-8 | Sequence 8, Appli |
| 8 | 31 | 68.9 | 441 | 1 | US-08-483-232-8 | Sequence 8, Appli |
| 9 | 31 | 68.9 | 441 | 1 | US-08-483-140-8 | Sequence 8, Appli |
| 10 | 31 | 68.9 | 441 | 2 | US-08-485-938A-8 | Sequence 8, Appli |
| 11 | 31 | 68.9 | 441 | 2 | US-08-910-041-8 | Sequence 8, Appli |
| 12 | 31 | 68.9 | 441 | 3 | US-09-328-474-8 | Sequence 8, Appli |
| 13 | 31 | 68.9 | 441 | 3 | US-09-100-546-8 | Sequence 8, Appli |
| 14 | 31 | 68.9 | 441 | 4 | US-09-010-715-8 | Sequence 8, Appli |
| 15 | 31 | 68.9 | 441 | 4 | US-09-577-758-8 | Sequence 8, Appli |
| 16 | 31 | 68.9 | 444 | 1 | US-08-483-140-28 | Sequence 28, Appl |
| 17 | 31 | 68.9 | 444 | 2 | US-08-485-938A-32 | Sequence 32, Appl |
| 18 | 31 | 68.9 | 902 | 1 | US-08-701-846-2 | Sequence 2, Appli |
| 19 | 31 | 68.9 | 2710 | 1 | US-08-480-604A-6 | Sequence 6, Appli |
| 20 | 31 | 68.9 | 2710 | 2 | US-08-405-496A-6 | Sequence 6, Appli |
| 21 | 31 | 68.9 | 2710 | 4 | US-08-915-136-6 | Sequence 6, Appli |
| 22 | 30 | 66.7 | 84 | 4 | US-09-227-357-597 | Sequence 597, App |
| 23 | 30 | 66.7 | 572 | 6 | 5200183-5 | Patent No. 5200183 |
| 24 | 30 | 66.7 | 602 | 3 | US-08-446-100-1 | Sequence 1, Appli |
| 25 | 30 | 66.7 | 602 | 3 | US-08-446-100-2 | Sequence 2, Appli |
| 26 | 30 | 66.7 | 602 | 3 | US-08-446-100-3 | Sequence 3, Appli |
| 27 | 30 | 66.7 | 602 | 3 | US-08-446-100-4 | Sequence 4, Appli |
| 28 | 30 | 66.7 | 602 | 3 | US-08-446-100-5 | Sequence 5, Appli |
| 29 | 30 | 66.7 | 602 | 3 | US-08-446-100-6 | Sequence 6, Appli |
| 30 | 30 | 66.7 | 602 | 3 | US-08-446-100-7 | Sequence 7, Appli |
| 31 | 30 | 66.7 | 602 | 3 | US-08-446-100-8 | Sequence 8, Appli |
| 32 | 30 | 66.7 | 602 | 3 | US-08-446-100-9 | Sequence 9, Appli |
| 33 | 30 | 66.7 | 602 | 3 | US-08-446-100-10 | Sequence 10, Appl |
| 34 | 30 | 66.7 | 602 | 3 | US-08-446-100-11 | Sequence 11, Appl |
| 35 | 30 | 66.7 | 602 | 3 | US-08-446-100-12 | Sequence 12, Appl |
| 36 | 30 | 66.7 | 602 | 3 | US-08-446-100-13 | Sequence 13, Appl |
| 37 | 30 | 66.7 | 602 | 3 | US-08-446-100-14 | Sequence 14, Appl |
| 38 | 30 | 66.7 | 602 | 3 | US-08-446-100-15 | Sequence 15, Appl |
| 39 | 30 | 66.7 | 602 | 3 | US-08-446-100-16 | Sequence 16, Appl |
| 40 | 30 | 66.7 | 602 | 3 | US-08-446-100-17 | Sequence 17, Appl |
| 41 | 30 | 66.7 | 602 | 3 | US-08-446-100-18 | Sequence 18, Appl |
| 42 | 30 | 66.7 | 602 | 3 | US-08-446-100-24 | Sequence 24, Appl |
| 43 | 30 | 66.7 | 602 | 4 | US-09-334-489-3 | Sequence 3, Appli |
| 44 | 30 | 66.7 | 602 | 4 | US-09-334-489-4 | Sequence 4, Appli |
| 45 | 30 | 66.7 | 602 | 6 | 5215909-11 | Patent No. 5215909 |

ALIGNMENTS

```
RESULT   1
US-08-193-114B-1
; Sequence 1, Application US/08193114B
; Patent No. 5472945
;  GENERAL INFORMATION:
;    APPLICANT:  Schmaier, Alvin H.
;    APPLICANT:  Jiang, Yongping
;    TITLE OF INVENTION:  Modulation of Blood
;    TITLE OF INVENTION:  Pressure and Inhibition of Platelet Activation
;    TITLE OF INVENTION:  with Kininogen Fragment
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seidel, Gonda, Lavorgna &
;      ADDRESSEE:  Monaco, P.C.
;      STREET:  1800 Two Penn Center Plaza
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/193,114B
;      FILING DATE:  9 February 1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  U.S. Application
;      APPLICATION NUMBER:  Serial No. 5472945 07/744,545
;      FILING DATE:  13 August 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-137 CI1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  No. 5472945e
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  117 amino acids
;      TYPE:  peptide
;      TOPOLOGY:  linear
US-08-193-114B-1


  Query Match             100.0%;  Score 45;  DB 1;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 0.16;
```

```
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NNATFYFK 8
          ||||||||
Db     30 NNATFYFK 37


RESULT   2
PCT-US92-06809-1
; Sequence 1, Application PC/TUS9206809
;  GENERAL INFORMATION:
;    APPLICANT:  Schmaier, Alvin H.
;    APPLICANT:  Jiang, Yongping
;    TITLE OF INVENTION:  Modulation of Blood
;    TITLE OF INVENTION:  Pressure by Altering Bradykinin Levels
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of the
;      ADDRESSEE:  Commonwealth System of Higher Education
;      STREET:  406 University Services
;      STREET:  Building
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/06809
;      FILING DATE:  19910813
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  U.S. Application
;      APPLICATION NUMBER:  Serial No.  744,545
;      FILING DATE:  13 August 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-137
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  117 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
PCT-US92-06809-1


  Query Match            100.0%;  Score 45;  DB 5;  Length 117;
  Best Local Similarity  100.0%;  Pred. No. 0.16;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NNATFYFK 8
          ||||||||
Db     30 NNATFYFK 37


RESULT   3
US-08-676-242-15
; Sequence 15, Application US/08676242C
; Patent No. 6143719
; GENERAL INFORMATION:
;  APPLICANT: The Regents of the University of Michigan
;  APPLICANT: Schmaier, Alvin H.
;  APPLICANT: Hasan, Ahmed A.K.
;  TITLE OF INVENTION: Bradykinin Analogs As Selective Thrombin Inhibitors
;  FILE REFERENCE: 8820-2 US
;  CURRENT APPLICATION NUMBER: US/08/676,242C
;  CURRENT FILING DATE: 2000-07-16
;  EARLIER APPLICATION NUMBER: 60/000,096
;  EARLIER FILING DATE: 1995-06-09
;  EARLIER APPLICATION NUMBER: PCT/US96/09940
;  EARLIER FILING DATE: 1996-06-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 26
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Bradykinin
;   OTHER INFORMATION: analog
US-08-676-242-15


  Query Match            86.7%;  Score 39;  DB 4;  Length 26;
  Best Local Similarity  100.0%;  Pred. No. 0.43;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 NATFYFK 8
          |||||||
Db      1 NATFYFK 7


RESULT   4
US-08-483-140-30
; Sequence 30, Application US/08483140
; Patent No. 5698403
;  GENERAL INFORMATION:
;    APPLICANT:  ICOS Corporation
;    TITLE OF INVENTION:  Platelet-Activating Factor Acetyl
;    TITLE OF INVENTION:  Hydrolase
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gertein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,140
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/318,905
;      FILING DATE:  6-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  6-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5698403and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  32781
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  193 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-483-140-30
```

Query Match 68.9%; Score 31; DB 1; Length 193;
Best Local Similarity 71.4%; Pred. No. 89;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy      2 NATFYFK 8
          :||:|||
Db    143 SATYYFK 149
```

```
RESULT   5
US-08-485-938A-36
; Sequence 36, Application US/08485938A
; Patent No. 5847088
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor
;    TITLE OF INVENTION:  Acetylhydrolase
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,938A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/318,905
;      FILING DATE:  06-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  06-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5847088and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/32792
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  193 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-485-938A-36
```

Query Match 68.9%; Score 31; DB 2; Length 193;
Best Local Similarity 71.4%; Pred. No. 89;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy      2 NATFYFK 8
          :||:|||
Db    143 SATYYFK 149
```

```
RESULT   6
US-08-470-187-8
; Sequence 8, Application US/08470187
; Patent No. 5532152
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine E.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor Acetyl
;    TITLE OF INVENTION:  Hydrolase
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gertein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,187
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5532152and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  31672
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  441 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-470-187-8
```

Query Match 68.9%; Score 31; DB 1; Length 441;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy      2 NATFYFK 8
          :||:|||
Db    185 SATYYFK 191
```

```
RESULT   7
US-08-318-905-8
; Sequence 8, Application US/08318905
; Patent No. 5641669
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor Acetyl
;    TITLE OF INVENTION:  Hydrolase
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  Marshall, O'Toole, Gertein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/318,905
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  6-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5641669and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  32205
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  441 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-318-905-8


  Query Match             68.9%;  Score 31;  DB 1;  Length 441;
  Best Local Similarity   71.4%;  Pred. No. 2e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 NATFYFK 8
           :||:|||
Db     185 SATYYFK 191


RESULT   8
US-08-483-232-8
; Sequence 8, Application US/08483232
; Patent No. 5656431
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor
;    TITLE OF INVENTION:  Acetylhydrolase
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,232
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/318,905
;      FILING DATE:  06-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  06-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5656431and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/32689
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  441 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-483-232-8


  Query Match             68.9%;  Score 31;  DB 1;  Length 441;
  Best Local Similarity   71.4%;  Pred. No. 2e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 NATFYFK 8
           :||:|||
Db     185 SATYYFK 191


RESULT   9
US-08-483-140-8
; Sequence 8, Application US/08483140
; Patent No. 5698403
;  GENERAL INFORMATION:
;    APPLICANT:  ICOS Corporation
;    TITLE OF INVENTION:  Platelet-Activating Factor Acetyl
;    TITLE OF INVENTION:  Hydrolase
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gertein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,140
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/318,905
;      FILING DATE:  6-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  6-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5698403and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  32781
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  8:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  441 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-483-140-8


  Query Match             68.9%;  Score 31;  DB 1;  Length 441;
  Best Local Similarity   71.4%;  Pred. No. 2e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 NATFYFK 8
            :||:|||
Db    185 SATYYFK 191


RESULT  10
US-08-485-938A-8
; Sequence 8, Application US/08485938A
; Patent No. 5847088
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor
;    TITLE OF INVENTION:  Acetylhydrolase
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,938A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/318,905
;      FILING DATE:  06-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  06-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5847088and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/32792
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  441 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-485-938A-8


  Query Match             68.9%;  Score 31;  DB 2;  Length 441;
  Best Local Similarity   71.4%;  Pred. No. 2e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 NATFYFK 8
            :||:|||
Db    185 SATYYFK 191


RESULT  11
US-08-910-041-8
; Sequence 8, Application US/08910041
; Patent No. 5977308
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor
;    TITLE OF INVENTION:  Acetylhydrolase
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/910,041
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/483,232
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/318,905
;      FILING DATE:  06-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  06-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27866/34026
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  441 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-910-041-8


  Query Match             68.9%;  Score 31;  DB 2;  Length 441;
  Best Local Similarity   71.4%;  Pred. No. 2e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 NATFYFK 8
            :||:|||
Db    185 SATYYFK 191


RESULT  12
US-09-328-474-8
```

; Sequence 8, Application US/09328474
; Patent No. 6045794
; GENERAL INFORMATION:
; APPLICANT: Cousens, Lawrence S.
; APPLICANT: Eberhardt, Christine D.
; APPLICANT: Gray, Patrick W.
; APPLICANT: Le Trong, Hai
; APPLICANT: Tjoelker, Larry W.
; APPLICANT: Wilder, Cheryl L.
; TITLE OF INVENTION: Platelet-Activating Factor
; TITLE OF INVENTION: Acetylhydrolase
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/328,474
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/483,232
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/318,905
; FILING DATE: 06-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/133,803
; FILING DATE: 06-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3658
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-328-474-8

Query Match 68.9%; Score 31; DB 3; Length 441;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NATFYFK 8
:||:|||
Db 185 SATYYFK 191

RESULT 13
US-09-100-546-8
; Sequence 8, Application US/09100546
; Patent No. 6099836
; GENERAL INFORMATION:
; APPLICANT: Cousens, Lawrence S.
; APPLICANT: Eberhardt, Christine D.
; APPLICANT: Gray, Patrick W.
; APPLICANT: Le Trong, Hai
; APPLICANT: Tjoelker, Larry W.
; APPLICANT: Wilder, Cheryl L.
; TITLE OF INVENTION: Platelet-Activating Factor
; TITLE OF INVENTION: Acetylhydrolase
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,546
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/010,715
; FILING DATE:
; APPLICATION NUMBER: US 08/318,905
; FILING DATE: 06-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/133,803
; FILING DATE: 06-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6099836and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/32793
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3658
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-100-546-8

Query Match 68.9%; Score 31; DB 3; Length 441;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NATFYFK 8
:||:|||
Db 185 SATYYFK 191

RESULT 14
US-09-010-715-8
; Sequence 8, Application US/09010715
; Patent No. 6146625
; GENERAL INFORMATION:
; APPLICANT: Cousens, Lawrence S.
; APPLICANT: Eberhardt, Christine D.
; APPLICANT: Gray, Patrick W.
; APPLICANT: Le Trong, Hai
; APPLICANT: Tjoelker, Larry W.
; APPLICANT: Wilder, Cheryl L.
; TITLE OF INVENTION: Platelet-Activating Factor
; TITLE OF INVENTION: Acetylhydrolase
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois

```
;       COUNTRY:  United States of America
;       ZIP:  60606-6402
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/010,715
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/318,905
;       FILING DATE:  06-OCT-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/133,803
;       FILING DATE:  06-OCT-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  No. 6146625and, Greta E.
;       REGISTRATION NUMBER:  35,302
;       REFERENCE/DOCKET NUMBER:  27866/32793
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 474-6300
;       TELEFAX:  (312) 474-0448
;       TELEX:  25-3658
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  441 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-09-010-715-8


  Query Match             68.9%;  Score 31;  DB 4;  Length 441;
  Best Local Similarity   71.4%;  Pred. No. 2e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 NATFYFK 8
           :||:|||
Db     185 SATYYFK 191


RESULT  15
US-09-577-758-8
; Sequence 8, Application US/09577758
; Patent No. 6203790
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor
;    TITLE OF INVENTION:  Acetylhydrolase
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/577,758
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/010,715
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  06-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6203790and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/32793
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  441 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-577-758-8


  Query Match             68.9%;  Score 31;  DB 4;  Length 441;
  Best Local Similarity   71.4%;  Pred. No. 2e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 NATFYFK 8
           :||:|||
Db     185 SATYYFK 191


Search completed: July  1, 2002, 16:18:00
Job time: 44 sec
```

GenCore version 4.5


OM protein - protein search, using sw model

Run on: July 1, 2002, 16:20:37 ; Search time 46.58 Seconds
(without alignments)
24.755 Million cell updates/sec

Title: US-09-461-061A-3
Perfect score: 56
Sequence: 1 IDNVKKARVQVV 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 56 | 100.0 | 427 | 1 | KGHUL1 | kininogen, LMW pre |
| 2 | 56 | 100.0 | 644 | 1 | KGHUH1 | kininogen, HMW pre |
| 3 | 44 | 78.6 | 434 | 1 | KGBOL2 | kininogen, LMW II |
| 4 | 44 | 78.6 | 436 | 1 | KGBOL1 | kininogen, LMW I p |
| 5 | 44 | 78.6 | 619 | 1 | KGBOH2 | kininogen, HMW II |
| 6 | 44 | 78.6 | 621 | 1 | KGBOH1 | kininogen, HMW I p |
| 7 | 40 | 71.4 | 540 | 2 | A71610 | HSP60 fold T-compl |
| 8 | 39 | 69.6 | 309 | 2 | G71103 | hypothetical prote |
| 9 | 39 | 69.6 | 768 | 2 | AH1085 | hypothetical prote |
| 10 | 38 | 67.9 | 423 | 1 | KGRTM | major acute phase |
| 11 | 38 | 67.9 | 430 | 1 | KGRTT1 | T-kininogen I prec |
| 12 | 38 | 67.9 | 430 | 2 | A23897 | major acute phase |
| 13 | 38 | 67.9 | 430 | 2 | B28055 | T-kininogen, LMW I |
| 14 | 38 | 67.9 | 433 | 2 | A28055 | K-kininogen, LMW I |
| 15 | 38 | 67.9 | 639 | 2 | A25486 | kininogen, HMW I p |
| 16 | 36 | 64.3 | 322 | 1 | RGBYC3 | regulatory protein |
| 17 | 36 | 64.3 | 585 | 2 | D57150 | hydrogenase (EC 1. |
| 18 | 35 | 62.5 | 222 | 2 | AI1190 | sortase homolog lm |
| 19 | 35 | 62.5 | 222 | 2 | AI1548 | sortase homolog li |
| 20 | 35 | 62.5 | 264 | 2 | F72216 | hypothetical prote |
| 21 | 35 | 62.5 | 334 | 2 | C95272 | conserved hypothet |
| 22 | 35 | 62.5 | 367 | 2 | S14619 | branched-chain ami |
| 23 | 35 | 62.5 | 421 | 2 | A99409 | conserved hypothet |
| 24 | 35 | 62.5 | 479 | 2 | C75513 | probable ferredoxi |
| 25 | 35 | 62.5 | 488 | 2 | T33626 | hypothetical prote |
| 26 | 35 | 62.5 | 552 | 2 | C87388 | hypothetical prote |
| 27 | 35 | 62.5 | 562 | 2 | S57083 | t-complex-type mol |
| 28 | 35 | 62.5 | 759 | 2 | AF1449 | hypothetical prote |
| 29 | 35 | 62.5 | 818 | 2 | F89819 | endopeptidase [imp |
| 30 | 35 | 62.5 | 881 | 2 | T49279 | hypothetical prote |
| 31 | 35 | 62.5 | 2342 | 2 | T18200 | fatty-acid synthas |
| 32 | 34 | 60.7 | 158 | 2 | E71709 | ATP synthase B cha |
| 33 | 34 | 60.7 | 203 | 2 | T10885 | auxin-induced prot |
| 34 | 34 | 60.7 | 227 | 2 | T37535 | vacuolar atp synth |
| 35 | 34 | 60.7 | 246 | 2 | S23538 | hypothetical prote |
| 36 | 34 | 60.7 | 249 | 2 | S74373 | heat shock protein |
| 37 | 34 | 60.7 | 250 | 2 | G75217 | glycerophosphodies |
| 38 | 34 | 60.7 | 262 | 2 | A64142 | hypothetical prote |
| 39 | 34 | 60.7 | 264 | 2 | H87293 | hypothetical prote |
| 40 | 34 | 60.7 | 363 | 1 | B69757 | toxic anion resist |
| 41 | 34 | 60.7 | 380 | 2 | H71191 | hypothetical prote |
| 42 | 34 | 60.7 | 389 | 2 | B44972 | paramyosin - nemat |
| 43 | 34 | 60.7 | 400 | 2 | A47094 | aspartate transami |
| 44 | 34 | 60.7 | 424 | 2 | C64362 | aconitate hydratas |
| 45 | 34 | 60.7 | 566 | 1 | F64329 | arginine--tRNA lig |

ALIGNMENTS

RESULT 1
KGHUL1
kininogen, LMW precursor [validated] - human
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Homo sapiens (man)
C;Date: 06-Jul-1982 #sequence_revision 27-Nov-1985 #text_change 08-Dec-2000
C;Accession: A01280; B25276; A27900; A27699; A31905; A34030
R;Ohkubo, I.; Kurachi, K.; Takasawa, T.; Shiokawa, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A;Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its ide
A;Reference number: A90490; MUID:85122621
A;Accession: A01280
A;Molecule type: mRNA
A;Residues: 1-427 <OHK>
A;Cross-references: GB:K02566; NID:g177889; PIDN:AAA35497.1; PID:g177890
R;Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A;Title: Cloning and sequence analysis of cDNAs for human high molecular weight and l
A;Reference number: A92544; MUID:85234582
A;Accession: B25276
A;Molecule type: mRNA
A;Residues: 1-427 <TAK>
A;Cross-references: GB:M11437; NID:g186751; PIDN:AAB59551.1; PID:g386853
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Rauth, G.; Mueller-Esterl, W.
in Kinins IV, part A, Greenbaum, L.M., and Margolius, H.S., eds., pp.91-95, Plenum, N
A;Title: Amino acid sequence of the light chain of human low molecular mass kininogen
A;Reference number: A27900
A;Accession: A27900
A;Molecule type: protein
A;Residues: 390-427 <LOT>
R;Mindroiu, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
Biochem. Biophys. Res. Commun. 152, 519-526, 1988
A;Title: A new kinin moiety in human plasma kininogens.
A;Reference number: A27699; MUID:88209021
A;Accession: A27699
A;Molecule type: protein
A;Residues: 380-389 <MIN>
R;Maeda, H.; Matsumura, Y.; Kato, H.
J. Biol. Chem. 263, 16051-16054, 1988
A;Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic f
A;Reference number: A31905; MUID:89034061
A;Accession: A31905
A;Molecule type: protein
A;Residues: 381-389 <MAE>
R;Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
Biochem. Biophys. Res. Commun. 150, 511-516, 1988
A;Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human p
A;Reference number: A34030; MUID:88106632
A;Accession: A34030
A;Molecule type: protein
A;Residues: 380-389 <SAS>

R;Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 260, 8610-8617, 1985
A;Title: Structural organization of the human kininogen gene and a model for its evoluti
A;Reference number: A92545; MUID:85234583
A;Contents: annotation; gene organization
R;Pierce, J.V.
Fed. Proc. 27, 52-57, 1968
A;Title: Structural features of plasma kinins and kininogens.
A;Reference number: A91455; MUID:90255622
A;Contents: annotation; bradykinin
C;Comment: The LMW kininogen precursor is produced from the same gene as the HMW form (s
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Genetics:
A;Gene: GDB:KNG
A;Cross-references: GDB:125256; OMIM:228960
A;Map position: 3q27-3q27
A;Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3; 401/3
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; glyc
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-427/Product: LMW prokininogen (kininogen I) #status predicted <MAT>
F;19-389,390-427/Product: LMW kininogen II #status predicted <MAT2>
F;19-379/Product: LMW kininogen heavy chain #status predicted <HCH>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>
F;380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
F;390-427/Product: LMW kininogen light chain #status experimental <LCH>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;28-407,83-94,107-126,142-145,206-218,229-248,264-267,328-340,351-370/Disulfide bonds:
F;48,169,205,294/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
F;383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F;389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;401/Binding site: carbohydrate (Thr) (covalent) #status absent

Query Match 100.0%; Score 56; DB 1; Length 427;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 IDNVKKARVQVV 12
           ||||||||||||
Db     301 IDNVKKARVQVV 312
```

RESULT 2
KGHUH1
kininogen, HMW precursor [validated] - human
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen; prokininogen
N;Contains: bradykinin (kallidin I); HMW kininogen I; HMW kininogen II; low molecular we
C;Species: Homo sapiens (man)
C;Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 08-Dec-2000
C;Accession: A01279; A25276; S32422; A91153; A24871; A27899; A27699; A31905; A34030; S02
R;Ohkubo, I.; Kurachi, K.; Takasawa, T.; Shiokawa, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A;Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its identi
A;Reference number: A90490; MUID:85122621
A;Accession: A01279
A;Molecule type: mRNA
A;Residues: 1-389 <OHK>
A;Cross-references: GB:K02566; NID:g177889
R;Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A;Title: Cloning and sequence analysis of cDNAs for human high molecular weight and low
A;Reference number: A92544; MUID:85234582
A;Accession: A25276
A;Molecule type: mRNA
A;Residues: 1-592,'I',594-644 <TAK>
A;Cross-references: GB:M11437; NID:g186751; PIDN:AAB59550.1; PID:g386852
R;Auerswald, E.A.; Roessler, D.; Mentele, R.; Assfalg-Machleidt, I.
FEBS Lett. 321, 93-97, 1993
A;Title: Cloning, expression and characterization of human kininogen domain 3.
A;Reference number: S32422; MUID:93223854
A;Accession: S32422
A;Molecule type: mRNA
A;Residues: 'ANSM',253-377 <AUE>
A;Note: differences are due to known cloning artifacts
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A;Title: The amino acid sequence of the light chain of human high-molecular-mass kini
A;Reference number: A91153; MUID:86030270
A;Accession: A91153
A;Molecule type: protein
A;Residues: 379-644 <LOT>
A;Note: the bradykinin sequence preceding the light chain sequence was not determined
R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
Eur. J. Biochem. 154, 471-478, 1986
A;Title: Completion of the primary structure of human high-molecular-mass kininogen.
A;Reference number: A24871; MUID:86108361
A;Accession: A24871
A;Molecule type: protein
A;Residues: 'Z',20-380 <KEL1>
R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
in Kinins IV, Greenbaum, L.M., and Margolius, H.S., ed., pp.85-89, Plenum Press, New
A;Title: Amino acid sequence of the light chain of human high molecular mass kininoge
A;Reference number: A27899
A;Accession: A27899
A;Molecule type: protein
A;Residues: 379-389,'K',390-407,'Q',409-644 <KEL2>
R;Mindroiu, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
Biochem. Biophys. Res. Commun. 152, 519-526, 1988
A;Title: A new kinin moiety in human plasma kininogens.
A;Reference number: A27699; MUID:88209021
A;Accession: A27699
A;Molecule type: protein
A;Residues: 380-389 <MIN>
R;Maeda, H.; Matsumura, Y.; Kato, H.
J. Biol. Chem. 263, 16051-16054, 1988
A;Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic f
A;Reference number: A31905; MUID:89034061
A;Accession: A31905
A;Molecule type: protein
A;Residues: 381-389 <MAE>
R;Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
Biochem. Biophys. Res. Commun. 150, 511-516, 1988
A;Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human p
A;Reference number: A34030; MUID:88106632
A;Accession: A34030
A;Molecule type: protein
A;Residues: 380-389 <SAS>
R;Lenarcic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
Biol. Chem. Hoppe-Seyler 369, 257-261, 1988
A;Title: Human cathepsin B and cysteine proteinase inhibitors (CPIs) in inflammatory
A;Reference number: S02482; MUID:89076517
A;Accession: S02482
A;Molecule type: protein
A;Residues: 1-19;189-192;310-314;381-389 <LEN1>
R;Kato, H.; Matsumura, Y.; Maeda, H.
FEBS Lett. 232, 252-254, 1988
A;Title: Isolation and identification of hydroxyproline analogues of bradykinin in hu
A;Reference number: A61495; MUID:88211869
A;Accession: A61495
A;Molecule type: protein
A;Residues: 380-389 <KAT1>
A;Experimental source: urine
A;Note: this peptide had Pro-383 modified to 4-hydroxyproline
A;Accession: B61495
A;Molecule type: protein
A;Residues: 381-389 <KAT2>
A;Experimental source: urine
A;Note: this peptide had Pro-383 modified to 4-hydroxyproline
A;Accession: C61495

A;Molecule type: protein
A;Residues: 380-389 <KAT3>
R;Lenarcic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
FEBS Lett. 280, 211-215, 1991
A;Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
A;Reference number: S14303; MUID:91192133
A;Accession: S14447
A;Molecule type: protein
A;Residues: 264-359,'N',361-375 <LEN2>
R;Little, S.S.; Johnson, D.A.
Biochem. J. 307, 341-346, 1995
A;Title: Human mast cell tryptase isoforms: separation and examination of substrate-spec
A;Reference number: S55239; MUID:95251593
A;Accession: S55239
A;Molecule type: protein
A;Residues: 450-452,'X',454,'X',456 <LIT>
R;Straczek, J.; Maachi, F.; le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabet, P.; Bellevil
FEBS Lett. 373, 207-211, 1995
A;Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like
A;Reference number: S68059; MUID:96033974
A;Accession: S68059
A;Molecule type: protein
A;Residues: 431-434 <STR>
R;Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 260, 8610-8617, 1985
A;Title: Structural organization of the human kininogen gene and a model for its evoluti
A;Reference number: A92545; MUID:85234583
A;Contents: annotation; gene organization
R;Pierce, J.V.
Fed. Proc. 27, 52-57, 1968
A;Title: Structural features of plasma kinins and kininogens.
A;Reference number: A91455; MUID:90255622
A;Contents: annotation; bradykinin
C;Comment: The HMW kininogen precursor and the LMW form are produced from the same gene
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Genetics:
A;Gene: GDB:KNG
A;Cross-references: GDB:125256; OMIM:228960
A;Map position: 3q27-3q27
A;Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>
F;19-379,390-644/Product: HMW kininogen II #status experimental <MAT2>
F;19-379/Domain: HMW kininogen heavy chain #status experimental <HCH>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>
F;380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
F;390-644/Domain: HMW kininogen light chain #status experimental <LCH>
F;421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
F;431-434/Product: low molecular weight growth promoting factor #status experimental <GF
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;28-614,83-94,107-126,142-145,206-218,229-248,264-267,328-340,351-370/Disulfide bonds:
F;48/Binding site: carbohydrate (Asn) (covalent) #status absent
F;169,205,294/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
F;383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F;389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;401,533,542,546,557,571,593,628/Binding site: carbohydrate (Thr) (covalent) #status ex
F;577/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 56; DB 1; Length 644;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 IDNVKKARVQVV 12
           ||||||||||||
Db     301 IDNVKKARVQVV 312
```

RESULT 3
KGBOL2
kininogen, LMW II precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 28-May-1999
C;Accession: A01284
R;Nawa, H.; Kitamura, N.; Hirose, T.; Asai, M.; Inayama, S.; Nakanishi, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 90-94, 1983
A;Title: Primary structures of bovine liver low molecular weight kininogen precursors
A;Reference number: A93984; MUID:83117859
A;Accession: A01284
A;Molecule type: mRNA
A;Residues: 1-434 <NAW>
A;Cross-references: GB:V00427; GB:J00011; NID:g489; PIDN:CAA23710.1; PID:g490
C;Comment: The LMW kininogen precursor is produced from the same gene as the HMW form
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; g
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-434/Product: LMW kininogen II #status predicted <MAT>
F;19-377/Product: LMW kininogen I heavy chain #status predicted <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;261-372/Domain: cystatin homology <CY3>
F;377-386/Product: lysyl-bradykinin (kallidin II) #status predicted <KBDY>
F;378-386/Product: bradykinin (kallidin I) #status predicted <BDY>
F;387-434/Product: LMW kininogen I light chain #status experimental <LCH>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F;27-404,82-93,106-125,141-144,205-217,228-247,261-264,325-337,348-367/Disulfide bond
F;47,87,168,169,197,204,280/Binding site: carbohydrate (Asn) (covalent) #status predi
F;376-377/Cleavage site: Met-Lys (kallikrein) #status predicted
F;380/Modified site: 4-hydroxyproline (Pro) #status predicted
F;386-387/Cleavage site: Arg-Ser (kallikrein) #status predicted

Query Match 78.6%; Score 44; DB 1; Length 434;
Best Local Similarity 83.3%; Pred. No. 2;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 IDNVKKARVQVV 12
           || |||| ||||
Db     298 IDTVKKATVQVV 309
```

RESULT 4
KGBOL1
kininogen, LMW I precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
C;Accession: A01283
R;Nawa, H.; Kitamura, N.; Hirose, T.; Asai, M.; Inayama, S.; Nakanishi, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 90-94, 1983
A;Title: Primary structures of bovine liver low molecular weight kininogen precursors
A;Reference number: A93984; MUID:83117859
A;Accession: A01283
A;Molecule type: mRNA
A;Residues: 1-436 <NAW>
A;Cross-references: GB:J00010; GB:V00426; NID:g163256; PIDN:AAA30604.1; PID:g163257
C;Comment: The LMW kininogen precursor is produced from the same gene as the HMW form
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology

C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; glyc
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-436/Product: LMW kininogen I #status predicted <MAT>
F;19-378/Product: LMW kininogen I heavy chain #status predicted <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;379-388/Product: lysyl-bradykinin (kallidin II) #status predicted <KBDY>
F;380-388/Product: bradykinin (kallidin I) #status predicted <BDY>
F;389-436/Product: LMW kininogen I light chain #status experimental <LCH>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;27-406,82-93,106-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bonds:
F;47,87,168,169,197,204/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;378-379/Cleavage site: Met-Lys (kallikrein) #status predicted
F;382/Modified site: 4-hydroxyproline (Pro) #status predicted
F;388-389/Cleavage site: Arg-Ser (kallikrein) #status predicted

Query Match 78.6%; Score 44; DB 1; Length 436;
Best Local Similarity 83.3%; Pred. No. 2.1;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IDNVKKARVQVV 12
|| |||| ||||
Db 300 IDTVKKATVQVV 311

RESULT 5
KGBOH2
kininogen, HMW II precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
C;Accession: A01282; A91923; A91941; A91938; B29559
R;Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
Nature 305, 545-549, 1983
A;Title: A single gene for bovine high molecular weight and low molecular weight kininog
A;Reference number: A93317; MUID:84014106
A;Accession: A01282
A;Molecule type: mRNA
A;Residues: 1-619 <KIT>
A;Cross-references: GB:V01492; GB:K01758; NID:g493; PIDN:CAA24736.1; PID:g494
R;Kato, H.; Nagasawa, S.; Suzuki, T.
J. Biochem. 67, 313-323, 1970
A;Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds and
A;Reference number: A91923; MUID:70180420
A;Accession: A91923
A;Molecule type: protein
A;Residues: 376-391 <KAT>
R;Han, Y.N.; Kato, H.; Iwanaga, S.; Suzuki, T.
J. Biochem. 79, 1201-1222, 1976
A;Title: Primary structure of bovine plasma high-molecular-weight kininogen. The amino a
A;Reference number: A91941; MUID:76260155
A;Accession: A91941
A;Molecule type: protein
A;Residues: 387-455 <HAN>
A;Note: 398-Pro, 401-Val, and 455-Lys were also found
R;Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
J. Biochem. 77, 55-68, 1975
A;Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Ami
A;Reference number: A91938; MUID:75170265
A;Accession: A91938
A;Molecule type: protein
A;Residues: 456-496 <HA2>
R;Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga,
J. Biol. Chem. 262, 2768-2779, 1987
A;Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of c
A;Reference number: A92627; MUID:87137530
A;Accession: B29559
A;Molecule type: protein
A;Residues: 'Z',20-104,'E',106-256,'XX',257-376 <SUE>
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A;Title: The amino acid sequence of the light chain of human high-molecular-mass kini
A;Reference number: A91153; MUID:86030270
A;Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
R;Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
Seikagaku 56, 808, 1984
A;Title: Disulfide bonds in bovine HMW kininogens.
A;Reference number: A94300
A;Contents: annotation; disulfide bonds
A;Note: article in Japanese
C;Comment: The HMW kininogen precursor is produced from the same gene as the LMW form
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of
C;Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is i
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-619/Product: HMW kininogen II #status predicted <MAT>
F;19-376/Product: HMW kininogen II heavy chain #status experimental <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;261-372/Domain: cystatin homology <CY3>
F;377-386/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;378-386/Product: bradykinin (kallidin I) #status experimental <BDY>
F;387-619/Product: HMW kininogen II light chain #status experimental <LCH>
F;418-488/Region: glycine/histidine/lysine-rich
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
F;27-589,82-93,106-125,141-144,205-217,228-247,261-264,325-337,348-367/Disulfide bond
F;47/Binding site: carbohydrate (Asn) (covalent) #status absent
F;87,168,169,204,280/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
F;197/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F;376-377/Cleavage site: Met-Lys (kallikrein) #status experimental
F;380/Modified site: 4-hydroxyproline (Pro) #status predicted
F;386-387/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;396,400,404,510/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;397,398,518,522,534,546,551,568/Binding site: carbohydrate (Thr) (covalent) #status
F;496-497/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match 78.6%; Score 44; DB 1; Length 619;
Best Local Similarity 83.3%; Pred. No. 2.9;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IDNVKKARVQVV 12
|| |||| ||||
Db 298 IDTVKKATVQVV 309

RESULT 6
KGBOH1
kininogen, HMW I precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
C;Accession: A01281; A91923; A91938; A29559
R;Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
Nature 305, 545-549, 1983
A;Title: A single gene for bovine high molecular weight and low molecular weight kini
A;Reference number: A93317; MUID:84014106
A;Accession: A01281
A;Molecule type: mRNA
A;Residues: 1-621 <KIT>
A;Cross-references: GB:V01491; GB:K01757; NID:g491; PIDN:CAA24735.1; PID:g492
R;Kato, H.; Nagasawa, S.; Suzuki, T.
J. Biochem. 67, 313-323, 1970
A;Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds a
A;Reference number: A91923; MUID:70180420
A;Accession: A91923
A;Molecule type: protein
A;Residues: 378-393 <KAT>

R;Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
J. Biochem. 77, 55-68, 1975
A;Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Ami
A;Reference number: A91938; MUID:75170265
A;Accession: A91938
A;Molecule type: protein
A;Residues: 458-498 <HAN>
R;Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga,
J. Biol. Chem. 262, 2768-2779, 1987
A;Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of c
A;Reference number: A92627; MUID:87137530
A;Accession: A29559
A;Molecule type: protein
A;Residues: 'Z',20-123,'I',125-127,'I',129-378 <SUE>
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A;Title: The amino acid sequence of the light chain of human high-molecular-mass kininog
A;Reference number: A91153; MUID:86030270
A;Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
R;Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
Seikagaku 56, 808, 1984
A;Title: Disulfide bonds in bovine HMW kininogens.
A;Reference number: A94300
A;Contents: annotation; disulfide bonds
A;Note: article in Japanese
C;Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-621/Product: HMW prokininogen I #status predicted <MAT>
F;19-379/Product: HMW kininogen I heavy chain #status experimental <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;379-388/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;380-388/Product: bradykinin (kallidin I) #status experimental <BDY>
F;389-621/Product: HMW kininogen I light chain #status experimental <LCH>
F;417-488/Region: glycine/histidine/lysine-rich
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;27-591,82-93,106-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bonds:
F;87,168,169,204/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
F;197/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F;378-379/Cleavage site: Met-Lys (kallikrein) #status experimental
F;382/Modified site: 4-hydroxyproline (Pro) #status predicted
F;388-389/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;398,406,512/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;399,400,520,524,536,548,553,570/Binding site: carbohydrate (Thr) (covalent) #status ex
F;498-499/Cleavage site: Arg-Thr (kallikrein) #status experimental

  Query Match 78.6%; Score 44; DB 1; Length 621;
  Best Local Similarity 83.3%; Pred. No. 2.9;
  Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IDNVKKARVQVV 12
|| |||| ||||
Db 300 IDTVKKATVQVV 311

RESULT 7
A71610
HSP60 fold T-complex protein 1 PFB0635w - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C;Accession: A71610
R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
.; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A;Reference number: A71600; MUID:99021743
A;Accession: A71610
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-540 <GAR>
A;Cross-references: GB:AE001407; GB:AE001362; NID:g3845234; PIDN:AAC71916.1; PID:g384
A;Experimental source: clone 3D7
C;Genetics:
A;Gene: PFB0635w
C;Superfamily: molecular chaperone t-complex-type

  Query Match 71.4%; Score 40; DB 2; Length 540;
  Best Local Similarity 66.7%; Pred. No. 14;
  Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IDNVKKARVQVV 12
||| ||| | |:
Db 277 IDNFKKANVDVI 288

RESULT 8
G71103
hypothetical protein PH1090 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 05-Nov-1999
C;Accession: G71103
R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Se
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophili
A;Reference number: A71000; MUID:98344137
A;Accession: G71103
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-309 <KAW>
A;Cross-references: GB:AP000004; NID:g3236131; PIDN:BAA30189.1; PID:d1031132; PID:g32
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBa
C;Genetics:
A;Gene: PH1090

  Query Match 69.6%; Score 39; DB 2; Length 309;
  Best Local Similarity 54.5%; Pred. No. 13;
  Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DNVKKARVQVV 12
:|:||||::|:
Db 46 ENLKKARIKVI 56

RESULT 9
AH1085
hypothetical protein lmo0087 [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AH1085
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1085
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-768 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC98302.1; PID:g16409446; GSPDB:GN00177
A;Experimental source: strain EGD-e

```
C;Genetics:
A;Gene: lmo0087


  Query Match             69.6%;  Score 39;  DB 2;  Length 768;
  Best Local Similarity   63.6%;  Pred. No. 31;
  Matches    7;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQV 11
           ||:||:||::|
Db     463 IDDVKRARIKV 473


RESULT  10
KGRTM
major acute phase alpha-1 protein precursor - rat (fragment)
N;Contains: bradykinin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 12-Apr-1996
C;Accession: A01285
R;Cole, T.; Inglis, A.S.; Roxburgh, C.M.; Howlett, G.J.; Schreiber, G.
FEBS Lett. 182, 57-61, 1985
A;Title: Major acute phase alphal-protein of the rat is homologous to bovine kininogen a
A;Reference number: A01285; MUID:85127561
A;Accession: A01285
A;Molecule type: mRNA
A;Residues: 1-423 <COL>
C;Comment: This plasma glycoprotein inhibits cysteine proteinases. During acute inflamma
r.
C;Superfamily: kininogen; cystatin homology
C;Keywords: bradykinin; cysteine proteinase inhibitor; duplication; glycoprotein; inflam
F;1-11/Domain: signal sequence (fragment) #status predicted <SIG>
F;12-423/Product: major acute phase alpha-1 protein #status predicted <MAT>
F;12-123/Domain: cystatin homology <CY1>
F;134-245/Domain: cystatin homology <CY2>
F;256-367/Domain: cystatin homology <CY3>
F;371-379/Product: bradykinin #status predicted <BDY>
F;12/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;161,197/Binding site: carbohydrate (Asn) (covalent) #status predicted


  Query Match             67.9%;  Score 38;  DB 1;  Length 423;
  Best Local Similarity   75.0%;  Pred. No. 26;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           || ||||  |||
Db     293 IDTVKKATSQVV 304


RESULT  11
KGRTT1
T-kininogen I precursor - rat
N;Alternate names: 73K protein; LMW kininogen T-I
N;Contains: bradykinin; T-kinin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 22-Jun-1999
C;Accession: A01286; D25486; A28526; PL0193; JQ0027; B25488; A28525; S68036
R;Furuto-Kato, S.; Matsumoto, A.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 12054-12059, 1985
A;Title: Primary structures of the mRNAs encoding the rat precursors for bradykinin and
nhibitor.
A;Reference number: A92496; MUID:86008264
A;Accession: A01286
A;Molecule type: mRNA
A;Residues: 1-430 <FUR>
A;Cross-references: GB:M11883; NID:g205084; PIDN:AAA41489.1; PID:g205085
R;Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A;Title: Differing expression patterns and evolution of the rat kininogen gene family.
A;Reference number: A92625; MUID:87137443
A;Accession: D25486
A;Molecule type: DNA
A;Residues: 375-430 <KIT>
R;Enjyoji, K.; Kato, H.; Hayashi, I.; Oh-ishi, S.; Iwanaga, S.
J. Biol. Chem. 263, 973-979, 1988
A;Title: Purification and characterization of rat T-kininogens isolated from plasma o
A;Reference number: A92729; MUID:88087226
A;Accession: A28526
A;Molecule type: protein
A;Residues: 'E',20-48;376-430 <ENJ>
R;Kanda, S.; Sugiyama, K.; Takahashi, M.; Shumiya, S.; Tomino, S.; Nagase, S.
Jpn. J. Cancer Res. 81, 63-68, 1990
A;Title: Identification of a protein increasing in serum of Nagase analbuminemic rats
A;Reference number: PL0193; MUID:90216390
A;Accession: PL0193
A;Molecule type: mRNA
A;Residues: 330-420,'R',422-429,'P' <KAN>
R;Anderson, K.P.; Croyle, M.L.; Lingrel, J.B.
Gene 81, 119-128, 1989
A;Title: Primary structure of a gene encoding rat T-kininogen.
A;Reference number: JQ0027; MUID:90034172
A;Accession: JQ0027
A;Molecule type: DNA
A;Residues: 1-60,'E',62-113,'R',115-165,'F',167-178,'TKI',182-211,'F',213-256,'S',258
A;Experimental source: strain Sprague-Dawley
R;Kageyama, R.; Kitamura, N.; Ohkubo, H.; Nakanishi, S.
J. Biol. Chem. 262, 2345-2351, 1987
A;Title: Differing utilization of homologous transcription initiation sites of rat K
A;Reference number: A25488; MUID:87137465
A;Accession: B25488
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-48 <KAG>
A;Cross-references: GB:M14356; NID:g205090; PIDN:AAA41492.1; PID:g205091
R;Enjyoji, K.; Kato, H.; Hayashi, I.; Oh-ishi, S.; Iwanaga, S.
J. Biol. Chem. 263, 965-972, 1988
A;Title: Purification and characterization of two kinds of low molecular weight kinin
A;Reference number: A28525; MUID:88087225
A;Accession: A28525
A;Molecule type: protein
A;Residues: 376-430 <EN2>
R;Sierra, F.; Walter, R.; Vautravers, P.; Guigoz, Y.
Arch. Biochem. Biophys. 322, 333-338, 1995
A;Title: Identification of several isoforms of T-kininogen expressed in the liver of
A;Reference number: S68034; MUID:96032652
A;Accession: S68036
A;Molecule type: mRNA
A;Residues: 340-430 <SIE>
A;Experimental source: clone pSG17
C;Comment: At least three types of LMW kininogen precursors are present in rat plasma
ceding bradykinin.
C;Comment: T-kininogens contain T-kinin (I-S-bradykinin), a novel kinin isolated afte
d of an Arg or Lys, it is probably not released from its precursor by either tissue o
C;Comment: The T-kininogens are produced in response to an inflammatory stimulant.
C;Genetics:
A;Introns: 65/3; 102/3; 130/1; 187/3; 223/2; 252/1; 309/3; 345/3; 374/3; 398/3
C;Superfamily: kininogen; cystatin homology
C;Keywords: acute phase; bradykinin; cysteine proteinase inhibitor; duplication; glyc
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-430/Product: T-kininogen I #status experimental <MAT>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;378-386/Product: bradykinin #status predicted <BDY>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
F;82,126,168,204,326/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;83-94,107-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bonds: #sta


  Query Match             67.9%;  Score 38;  DB 1;  Length 430;
  Best Local Similarity   75.0%;  Pred. No. 27;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARVQVV 12
           || ||||  |||
```

```
Db    300 IDTVKKATSQVV 311


RESULT  12
A23897
major acute phase alpha-1 protein (version 2) - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 20-Aug-1999
C;Accession: A23897; B23897
R;Anderson, K.P.; Heath, E.C.
J. Biol. Chem. 260, 12065-12071, 1985
A;Title: The relationship between rat major acute phase protein and the kininogens.
A;Reference number: A23897; MUID:86008266
A;Accession: A23897
A;Molecule type: protein
A;Residues: 1-14 <AND1>
A;Accession: B23897
A;Molecule type: mRNA
A;Residues: 5-430 <AND2>
A;Cross-references: GB:M11661; NID:g205307; PIDN:AAA41570.1; PID:g205308
A;Note: the authors translated the codon CTC for residue 410 as Arg, CTA for residue 415
C;Superfamily: kininogen; cystatin homology
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>


  Query Match             67.9%;  Score 38;  DB 2;  Length 430;
  Best Local Similarity   75.0%;  Pred. No. 27;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARVQVV 12
          || ||||  |||
Db    300 IDTVKKATSQVV 311


RESULT  13
B28055
T-kininogen, LMW II precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 12-Dec-1997
C;Accession: B28055; E25486; B28526; C28526
R;Furuto-Kato, S.; Matsumoto, A.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 12054-12059, 1985
A;Title: Primary structures of the mRNAs encoding the rat precursors for bradykinin and
nhibitor.
A;Reference number: A92496; MUID:86008264
A;Accession: B28055
A;Molecule type: mRNA
A;Residues: 1-430 <FUR>
R;Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A;Title: Differing expression patterns and evolution of the rat kininogen gene family.
A;Reference number: A92625; MUID:87137443
A;Accession: E25486
A;Molecule type: DNA
A;Residues: 375-430 <KIT>
R;Enjyoji, K.; Kato, H.; Hayashi, I.; Oh-ishi, S.; Iwanaga, S.
J. Biol. Chem. 263, 973-979, 1988
A;Title: Purification and characterization of rat T-kininogens isolated from plasma of a
A;Reference number: A92729; MUID:88087226
A;Accession: B28526
A;Molecule type: protein
A;Residues: 'E',20-25,'MD',28-48,376-430 <ENJ>
A;Accession: C28526
A;Molecule type: protein
A;Residues: 'E',20-48;376-388,'R',390-419,'ER',422-430 <EN2>
C;Superfamily: kininogen; cystatin homology
C;Keywords: glycoprotein; pyroglutamic acid
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-430/Product: T-kininogen, LMW II #status experimental <MAT>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
F;82,126,168,204,326/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;83-94,107-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bonds: #sta


  Query Match             67.9%;  Score 38;  DB 2;  Length 430;
  Best Local Similarity   75.0%;  Pred. No. 27;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARVQVV 12
          || ||||  |||
Db    300 IDTVKKATSQVV 311


RESULT  14
A28055
K-kininogen, LMW I precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 15-Nov-1996
C;Accession: A28055
R;Furuto-Kato, S.; Matsumoto, A.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 12054-12059, 1985
A;Title: Primary structures of the mRNAs encoding the rat precursors for bradykinin a
nhibitor.
A;Reference number: A92496; MUID:86008264
A;Accession: A28055
A;Molecule type: mRNA
A;Residues: 1-433 <FUR>
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-433/Product: K-kininogen, LMW I #status predicted <MAT>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>


  Query Match             67.9%;  Score 38;  DB 2;  Length 433;
  Best Local Similarity   75.0%;  Pred. No. 27;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARVQVV 12
          || ||||  |||
Db    301 IDTVKKATSQVV 312


RESULT  15
A25486
kininogen, HMW I precursor - rat
N;Contains: bradykinin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Nov-1996
C;Accession: A25486
R;Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A;Title: Differing expression patterns and evolution of the rat kininogen gene family
A;Reference number: A92625; MUID:87137443
A;Accession: A25486
A;Molecule type: mRNA
A;Residues: 1-639 <KIT>
A;Note: the authors translated the codon CAA for residue 347 as Asn
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-639/Product: kininogen, HMW I #status predicted <MAT>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>


  Query Match             67.9%;  Score 38;  DB 2;  Length 639;
  Best Local Similarity   75.0%;  Pred. No. 40;
```

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 IDNVKKARVQVV 12
           || ||||  |||
Db     301 IDTVKKATSQVV 312
```

Search completed: July 1, 2002, 16:20:38
Job time: 202 sec

OM protein - protein search, using sw model

Run on: July 1, 2002, 16:29:46 ; Search time 75.26 Seconds
(without alignments)
73.556 Million cell updates/sec

Title: US-09-461-061A-4
Perfect score: 161
Sequence: 1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 99 | 61.5 | 423 | 11 | P70517 | P70517 rattus norv |
| 2 | 98 | 60.9 | 430 | 11 | Q63581 | Q63581 rattus norv |
| 3 | 98 | 60.9 | 432 | 11 | Q91XK5 | Q91xk5 mus musculu |
| 4 | 61 | 37.9 | 167 | 11 | Q9QWL5 | Q9qwl5 mus musculu |
| 5 | 53 | 32.9 | 446 | 16 | Q99ZF4 | Q99zf4 streptococc |
| 6 | 52.5 | 32.6 | 268 | 3 | Q03984 | Q03984 saccharomyc |
| 7 | 52.5 | 32.6 | 498 | 3 | Q03983 | Q03983 saccharomyc |
| 8 | 52 | 32.3 | 537 | 3 | O13369 | O13369 ascobolus i |
| 9 | 51.5 | 32.0 | 803 | 17 | Q97UH8 | Q97uh8 sulfolobus |
| 10 | 51 | 31.7 | 139 | 5 | Q9TYY2 | Q9tyy2 caenorhabdi |
| 11 | 51 | 31.7 | 209 | 10 | Q43635 | Q43635 ricinus com |
| 12 | 51 | 31.7 | 523 | 16 | Q97GW8 | Q97gw8 clostridium |
| 13 | 51 | 31.7 | 928 | 17 | Q978G4 | Q978g4 thermoplasm |
| 14 | 50 | 31.1 | 503 | 10 | Q94IZ1 | Q94iz1 oryza sativ |
| 15 | 50 | 31.1 | 1690 | 3 | Q9P977 | Q9p977 candida alb |
| 16 | 50 | 31.1 | 1690 | 3 | Q9HFT8 | Q9hft8 candida alb |
| 17 | 50 | 31.1 | 1690 | 3 | Q9P411 | Q9p411 candida alb |
| 18 | 49.5 | 30.7 | 321 | 4 | Q9H6E8 | Q9h6e8 homo sapien |
| 19 | 49.5 | 30.7 | 692 | 4 | Q96LX0 | Q96lx0 homo sapien |
| 20 | 49.5 | 30.7 | 709 | 10 | Q9FVT7 | Q9fvt7 arabidopsis |
| 21 | 49 | 30.4 | 142 | 10 | O22479 | O22479 santalum al |
| 22 | 49 | 30.4 | 313 | 16 | Q9PP86 | Q9pp86 campylobact |
| 23 | 49 | 30.4 | 459 | 16 | Q98QI6 | Q98qi6 mycoplasma |
| 24 | 49 | 30.4 | 506 | 8 | Q9TLE3 | Q9tle3 diplycosia |
| 25 | 49 | 30.4 | 1034 | 16 | Q98PT0 | Q98pt0 mycoplasma |
| 26 | 49 | 30.4 | 5005 | 16 | Q9PPZ5 | Q9ppz5 ureaplasma |
| 27 | 48 | 29.8 | 167 | 4 | Q9UED4 | Q9ued4 homo sapien |
| 28 | 48 | 29.8 | 377 | 16 | O51327 | O51327 borrelia bu |
| 29 | 48 | 29.8 | 506 | 8 | O98788 | O98788 sphyrosperm |
| 30 | 48 | 29.8 | 506 | 8 | Q95GL1 | Q95gl1 diplycosia |
| 31 | 48 | 29.8 | 506 | 8 | Q95GK8 | Q95gk8 diplycosia |
| 32 | 48 | 29.8 | 507 | 8 | Q95GK9 | Q95gk9 diplycosia |
| 33 | 48 | 29.8 | 508 | 8 | O98786 | O98786 sphyrosperm |
| 34 | 48 | 29.8 | 517 | 12 | Q9YUS1 | Q9yus1 turkey aden |
| 35 | 48 | 29.8 | 556 | 4 | Q93000 | Q93000 homo sapien |
| 36 | 48 | 29.8 | 624 | 12 | Q9YVI1 | Q9yvi1 melanoplus |
| 37 | 48 | 29.8 | 680 | 4 | Q92498 | Q92498 homo sapien |
| 38 | 48 | 29.8 | 861 | 15 | Q9IUY9 | Q9iuy9 human immun |
| 39 | 48 | 29.8 | 906 | 4 | Q92998 | Q92998 homo sapien |
| 40 | 48 | 29.8 | 906 | 4 | Q92770 | Q92770 homo sapien |
| 41 | 48 | 29.8 | 970 | 4 | Q96FC9 | Q96fc9 homo sapien |
| 42 | 48 | 29.8 | 1231 | 10 | Q9SK98 | Q9sk98 arabidopsis |
| 43 | 47.5 | 29.5 | 107 | 5 | O45433 | O45433 caenorhabdi |
| 44 | 47.5 | 29.5 | 107 | 5 | Q94054 | Q94054 caenorhabdi |
| 45 | 47.5 | 29.5 | 144 | 5 | Q9VNT0 | Q9vnt0 drosophila |

ALIGNMENTS

RESULT 1
P70517
ID P70517 PRELIMINARY; PRT; 423 AA.
AC P70517;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MAJOR ACUTE PHASE ALPHA-1 PROTEIN PRECURSOR (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;.
RN [1]
RP SEQUENCE FROM N.A.
RA Cole T.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85149311; PubMed=2579644;
RA Cole T., Inglis A., Nagashima M., Schreiber G.;
RT "Major acute-phase alpha(1)-protein in the rat: Structure, molecular
RT cloning, and regulation of mRNA levels.";
RL Biochem. Biophys. Res. Commun. 126:719-724(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=85127561; PubMed=2578992;
RA Cole T., Inglis A.S., Roxburgh C.M., Howlett G.J., Schreiber G.;
RT "Major acute phase alpha1-protein of the rat is homologous to bovine
RT kininogen and contains the sequence for bradykinin: its synthesis is
RT regulated at the mRNA level.";
RL FEBS Lett. 182:57-61(1985).
DR EMBL; K02814; AAA41569.1; -.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR003243; Cystatin_C_M.
DR Pfam; PF00031; cystatin; 3.
DR ProDom; PD001231; Cystatin_C_M; 1.
DR SMART; SM00043; CY; 3.
DR PROSITE; PS00287; CYSTATIN; 2.
KW Signal.

```
FT   NON_TER       1      1
FT   SIGNAL       <1     11       POTENTIAL.
FT   CHAIN        12    423       POTENTIAL.
FT   CHAIN       371    379       POTENTIAL.
SQ   SEQUENCE   423 AA;  46905 MW;  F9E8BD3198547949 CRC64;


  Query Match             61.5%;  Score 99;  DB 11;  Length 423;
  Best Local Similarity   64.5%;  Pred. No. 9.2e-06;
  Matches   20;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
          | |:| :|||::|  ||||||  ||||  |||
Db    274 LGHSIARLNAQHNHIFYFKIDTVKKATSQVV 304


RESULT   2
Q63581
ID   Q63581      PRELIMINARY;      PRT;   430 AA.
AC   Q63581;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RAT T-KININOGEN (T-KG).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90034172; PubMed=2806908;
RA   Anderson K.P., Croyle M.L., Lingrel J.B.;
RT   "Primary structure of a gene encoding rat T-kininogen.";
RL   Gene 81:119-128(1989).
DR   EMBL; M29090; AAA42251.1; -.
DR   EMBL; M29083; AAA42251.1; JOINED.
DR   EMBL; M29084; AAA42251.1; JOINED.
DR   EMBL; M29091; AAA42251.1; JOINED.
DR   EMBL; M29085; AAA42251.1; JOINED.
DR   EMBL; M29086; AAA42251.1; JOINED.
DR   EMBL; M29087; AAA42251.1; JOINED.
DR   EMBL; M29088; AAA42251.1; JOINED.
DR   EMBL; M29089; AAA42251.1; JOINED.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   Pfam; PF00031; cystatin; 3.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
SQ   SEQUENCE   430 AA;  47618 MW;  45508DEF4BDC978C CRC64;


  Query Match             60.9%;  Score 98;  DB 11;  Length 430;
  Best Local Similarity   64.5%;  Pred. No. 1.3e-05;
  Matches   20;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
          | |:| :|||::|  ||||||  ||||  |||
Db    281 LGHSIAQLNAQHNHIFYFKIDTVKKATSQVV 311


RESULT   3
Q91XK5
ID   Q91XK5      PRELIMINARY;      PRT;   432 AA.
AC   Q91XK5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ADULT FEMALE PLACENTA CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE   CLONE:1600027I01, FULL INSERT SEQUENCE.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=PLACENTA;
RA   Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,
RA   Arakawa T., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA   Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA   Imotani K., Ishii Y., Itoh M., Izawa M., Kato H., Kawai J., Kojima Y.,
RA   Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Okazaki Y., Okido T.,
RA   Owa C., Saito H., Saito R., Sakai C., Sakai K., Sano H., Sasaki D.,
RA   Shibata K., Shibata Y., Shinagawa A., Shiraki T., Sogabe Y.,
RA   Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T., Tejima Y.,
RA   Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA   Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=PLACENTA;
RX   MEDLINE=21085660; PubMed=11217851;
RA   RIKEN FANTOM Consortium.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=PLACENTA;
RX   MEDLINE=99279253; PubMed=10349636;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=PLACENTA;
RX   MEDLINE=20499374; PubMed=11042159;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=PLACENTA;
RX   MEDLINE=20530913; PubMed=11076861;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
DR   EMBL; AK005547; BAB24115.1; -.
SQ   SEQUENCE   432 AA;  47898 MW;  91854EDA5284A16B CRC64;


  Query Match             60.9%;  Score 98;  DB 11;  Length 432;
  Best Local Similarity   64.5%;  Pred. No. 1.3e-05;
  Matches   20;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      2 LTHTITKLNAENNATFYFKIDNVKKARVQVV 32
          | |:| :|||||:  ||:||| ||||  |||
Db    281 LGHSIAQLNAENDHPFYYKIDTVKKATSQVV 311


RESULT   4
Q9QWL5
ID   Q9QWL5      PRELIMINARY;      PRT;   167 AA.
AC   Q9QWL5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
```

```
DE   MURINE CMAP (CYSTATIN F) (LEUKOCYSTATIN).
GN   MURINE CMAP OR CST7.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Morita M., Arakawa H., Yoshiuchi N.;
RT   "A novel cystatin-like metastasis associated gene.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AB015224; BAA34940.1; -.
DR   EMBL; AK004420; BAB23298.1; -.
DR   HSSP; P01034; 1G96.
DR   MGD; MGI:1298217; Cst7.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   Pfam; PF00031; cystatin; 1.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 1.
SQ   SEQUENCE   167 AA;  18847 MW;  61F776D8445095FE CRC64;


  Query Match             37.9%;  Score 61;  DB 11;  Length 167;
  Best Local Similarity   48.3%;  Pred. No. 0.73;
  Matches   14;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       4 HTITKLNAENNATFYFKIDNVKKARVQVV 32
           |:: | |   |  | ||  :| || ||||
Db      77 HSVEKFNNCTNDIFLFKESHVSKALVQVV 105


RESULT   5
Q99ZF4
ID   Q99ZF4      PRELIMINARY;      PRT;   446 AA.
AC   Q99ZF4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL PROTEIN SPY1252.
GN   SPY1252.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
RX   MEDLINE=21192684; PubMed=11296296;
RA   Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA   Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA   Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J.,
RA   Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT   "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR   EMBL; AE006565; AAK34107.1; -.
KW   Hypothetical protein; Complete proteome;
SQ   SEQUENCE   446 AA;  51511 MW;  F98FB7B071D668A3 CRC64;


  Query Match             32.9%;  Score 53;  DB 16;  Length 446;
  Best Local Similarity   39.3%;  Pred. No. 27;
  Matches   11;  Conservative    7;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAENNATFYFKIDNVKKAR 28
             :||: : ||  | :|||  |:  ::| |
Db     126 SLTNQLDKLALEKDATFQSKLATIEKER 153


RESULT   6
Q03984
ID   Q03984      PRELIMINARY;      PRT;   268 AA.
AC   Q03984;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL 31.1 KDA PROTEIN.
GN   YDR179W-A.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB972;
RA   Murphy L., Harris D.E.;
RL   Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB972;
RA   Barrell B., Rajandream M.A.;
RL   Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z46727; CAA86686.1; -.
DR   SGD; S0002587; YDR179W-A.
KW   Hypothetical protein.
SQ   SEQUENCE   268 AA;  31096 MW;  45C7F756F361C71D CRC64;


  Query Match             32.6%;  Score 52.5;  DB 3;  Length 268;
  Best Local Similarity   41.4%;  Pred. No. 19;
  Matches   12;  Conservative    5;  Mismatches   11;  Indels    1;  Gaps    1;

Qy       1 TLTHTITKLNAENNATFYFKIDNVKKARV 29
           || | |:|||| |    |   ||: : ::
Db     144 TLQHWISKLNALNR-VMYRTFDNIVQTKI 171


RESULT   7
Q03983
ID   Q03983      PRELIMINARY;      PRT;   498 AA.
AC   Q03983;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL 57.9 KDA PROTEIN.
GN   YDR179W-A.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=AB972;
RA   Murphy L., Harris D.E.;
RL   Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB972;
RA   Barrell B., Rajandream M.A.;
RL   Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z46727; CAA86685.1; -.
DR   SGD; S0002587; YDR179W-A.
KW   Hypothetical protein.
SQ   SEQUENCE   498 AA;  57916 MW;  60959427D7A230CD CRC64;


  Query Match             32.6%;  Score 52.5;  DB 3;  Length 498;
  Best Local Similarity   41.4%;  Pred. No. 36;
  Matches   12;  Conservative    5;  Mismatches   11;  Indels    1;  Gaps    1;

Qy       1 TLTHTITKLNAENNATFYFKIDNVKKARV 29
           || | |:|||| |    |   ||: : ::
Db     374 TLQHWISKLNALNR-VMYRTFDNIVQTKI 401


RESULT   8
O13369
ID   O13369      PRELIMINARY;      PRT;   537 AA.
AC   O13369;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PROBABLE DNA METHYLTRANSFERASE MASC1 (EC 2.1.1.-).
GN   MASC1.
OS   Ascobolus immersus.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Pezizomycetes;
OC   Pezizales; Ascobolaceae; Ascobolus.
OX   NCBI_TaxID=5191;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   STRAIN=RN42;
RX   MEDLINE=98004231; PubMed=9346245;
RA   Malagnac F., Wendel B., Goyon C., Faugeron G., Zickler D.,
RA   Rossignol J.-L., Noyer-Weidner M., Vollmayr P., Trautner T.A.,
RA   Walter J.;
RT   "A gene essential for de novo methylation and development in Ascobolus
RT   reveals a novel type of eukaryotic DNA methyltransferase structure.";
RL   Cell 91:281-290(1997).
CC   -!- FUNCTION: ESSENTIAL FOR DE NOVO METHYLATION AND SEXUAL
CC       DEVELOPMENT.
CC   -!- SIMILARITY: SOME, TO PROKARYOTIC MODIFICATION METHYLASES AND
CC       EUKARYOTIC DNA (CYTOSINE-5)-METHYLTRANSFERASES.
DR   EMBL; AF025475; AAC49849.1; -.
DR   REBASE; 3106; M.AimAI.
DR   InterPro; IPR001025; BAH.
DR   InterPro; IPR001525; C5_DNA_meth.
DR   Pfam; PF01426; BAH; 1.
DR   Pfam; PF00145; DNA_methylase; 2.
DR   PROSITE; PS00094; C5_MTASE_1; UNKNOWN_1.
DR   PROSITE; PS00095; C5_MTASE_2; UNKNOWN_1.
KW   Transferase; Methyltransferase; Developmental protein; DNA-binding.
FT   ACT_SITE    301    301       BY SIMILARITY.
SQ   SEQUENCE   537 AA;  61240 MW;  615EDA2F2453C174 CRC64;


  Query Match             32.3%;  Score 52;  DB 3;  Length 537;
  Best Local Similarity   45.8%;  Pred. No. 45;
  Matches   11;  Conservative    3;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       4 HTITKLNAENNATFYFKIDNVKKA 27
           |||   | |||   :| : |: ||
Db     308 HTIAGKNDENNEASFFAVVNLIKA 331


RESULT   9
Q97UH8
ID   Q97UH8      PRELIMINARY;      PRT;   803 AA.
AC   Q97UH8;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   HYPOTHETICAL PROTEIN SSO3039.
GN   SSO3039.
OS   Sulfolobus solfataricus.
OC   Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX   NCBI_TaxID=2287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35092 / DSM 1617 / P2;
RX   MEDLINE=21332296; PubMed=11427726;
RA   She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA   Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA   De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA   Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA   Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA   Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA   Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT   "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR   EMBL; AE006895; AAK43141.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   803 AA;  93417 MW;  2CAD55CDA1B009CD CRC64;


  Query Match             32.0%;  Score 51.5;  DB 17;  Length 803;
  Best Local Similarity   31.2%;  Pred. No. 81;
  Matches   10;  Conservative   10;  Mismatches   11;  Indels    1;  Gaps    1;

Qy       1 TLTHTITKLNAENNATFYFKIDN-VKKARVQV 31
           :|   | | |:| |  |:| ::|  :| ::::
Db     129 SLPVAIFKFNSEKNTRFFFGVNNPFEKGQIEI 160


RESULT  10
Q9TYY2
ID   Q9TYY2      PRELIMINARY;      PRT;   139 AA.
AC   Q9TYY2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   K08B4.6 PROTEIN (CYSTATIN).
GN   K08B4.6.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Rohlfing T., Murray J., Antonio B.;
RT   "The sequence of C. elegans cosmid K08B4.";
```

```
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Schierack P.S., Lucius R., Sonnenburg B., Hartmann S.A.E.;
RT   "Molecular cloning and characterization of Caenorhabditis elegans
RT   cystatin D.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF100663; AAC68983.1; -.
DR   EMBL; AJ310669; CAC33821.1; -.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
SQ   SEQUENCE   139 AA;  15091 MW;  796F1CD81166CEFE CRC64;


  Query Match             31.7%;  Score 51;  DB 5;  Length 139;
  Best Local Similarity   37.9%;  Pred. No. 15;
  Matches   11;  Conservative    8;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       4 HTITKLNAENNATFYFKIDNVKKARVQVV 32
           ::: ::|::||   |     | ||:|||
Db      39 NSVPEINSKNNGQNYMVPIKVVKAQVQVV 67


RESULT  11
Q43635
ID   Q43635      PRELIMINARY;      PRT;   209 AA.
AC   Q43635;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CYSTEINE PROTEINASE INHIBITOR.
OS   Ricinus communis (Castor bean).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Malpighiales; Euphorbiaceae; Ricinus.
OX   NCBI_TaxID=3988;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. CV. SANGUINEUS; TISSUE=COTYLEDON;
RA   Szederkenyi J., Schobert C.;
RT   "cDNA expressed in Ricinus cotyledons.";
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z49697; CAA89697.1; -.
SQ   SEQUENCE   209 AA;  23304 MW;  2E7BBF1D0D1DDFF1 CRC64;


  Query Match             31.7%;  Score 51;  DB 10;  Length 209;
  Best Local Similarity   26.7%;  Pred. No. 23;
  Matches    8;  Conservative    9;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFYFKIDNVKKARVQVV 32
           || :  :   :|: | :::  :  |: |||
Db     138 THAVNTIQQRSNSLFPYQLQEIVHAKAQVV 167


RESULT  12
Q97GW8
ID   Q97GW8      PRELIMINARY;      PRT;   523 AA.
AC   Q97GW8;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   SITE-SPECIFIC RECOMBINASE, DNA INVERTASE PIN HOMOLOG.
GN   CAC2247.
OS   Clostridium acetobutylicum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1488;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX   MEDLINE=21359325; PubMed=11466286;
RA   Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA   Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA   Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA   Bennett G.N., Koonin E.V., Smith D.R.;
RT   "Genome sequence and comparative analysis of the solvent-producing
RT   bacterium Clostridium acetobutylicum.";
RL   J. Bacteriol. 183:4823-4838(2001).
DR   EMBL; AE007725; AAK80204.1; -.
KW   Complete proteome.
SQ   SEQUENCE   523 AA;  62467 MW;  22DB00B060AF6102 CRC64;


  Query Match             31.7%;  Score 51;  DB 16;  Length 523;
  Best Local Similarity   46.7%;  Pred. No. 61;
  Matches   14;  Conservative    1;  Mismatches   11;  Indels    4;  Gaps    1;

Qy       2 LTHTITKLNAENNATFYFKIDNVKKARVQV 31
           ||  | | | |||     | ||  |  :||
Db     455 LTQEIEKFNVENN----LKCDNSLKNALQV 480


RESULT  13
Q978G4
ID   Q978G4      PRELIMINARY;      PRT;   928 AA.
AC   Q978G4;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TVG1502957 PROTEIN.
GN   TVG1502957.
OS   Thermoplasma volcanium.
OC   Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmaceae;
OC   Thermoplasma.
OX   NCBI_TaxID=50339;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GSS1 / DSM 4299 / JCM 9571;
RX   MEDLINE=20570466; PubMed=11121031;
RA   Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
RA   Kawashima-Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,
RA   Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;
RT   "Archaeal adaptation to higher temperatures revealed by genomic
RT   sequence of Thermoplasma volcanium.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).
DR   EMBL; AP000996; BAB60593.1; -.
DR   InterPro; IPR001087; Lipase_GDSL.
DR   InterPro; IPR000914; SBP_bac_5.
DR   Pfam; PF00496; SBP_bac_5; 1.
DR   PROSITE; PS01098; LIPASE_GDSL_SER; UNKNOWN_1.
KW   Complete proteome.
SQ   SEQUENCE   928 AA;  103142 MW;  6FB59B34A1B2B1BF CRC64;


  Query Match             31.7%;  Score 51;  DB 17;  Length 928;
  Best Local Similarity   52.6%;  Pred. No. 1.1e+02;
  Matches   10;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAENNATFYFK 20
           :|  ||  || || ||:|:
Db     448 ITQNITVNNATNNITFHFQ 466


RESULT  14
Q94IZ1
ID   Q94IZ1      PRELIMINARY;      PRT;   503 AA.
AC   Q94IZ1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
```

```
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE POLYGALACTURONASE-LIKE PROTEIN.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NIPPONBARE;
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
RT   clone:OSJNBa0025P13.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP003140; BAB55744.1; -.
SQ   SEQUENCE   503 AA;  52056 MW;  71DAABBC792DE6E4 CRC64;


  Query Match             31.1%;  Score 50;  DB 10;  Length 503;
  Best Local Similarity   36.0%;  Pred. No. 81;
  Matches    9;  Conservative    5;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       5 TITKLNAENNATFYFKIDNVKKARV 29
           |:  |  :|:  |:|: ||    ||
Db     251 TVRGLKVQNSPEFHFRFDNCNGVRV 275


RESULT  15
Q9P977
ID   Q9P977      PRELIMINARY;      PRT;  1690 AA.
AC   Q9P977;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CYR1.
GN   CYR1.
OS   Candida albicans (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5476;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ono N., Sudoh M.;
RT   "Candida albicans CYR1 gene.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB034965; BAA93553.1; -.
DR   InterPro; IPR001054; Guanylt_cyclase.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   InterPro; IPR001932; PP2C_domain.
DR   InterPro; IPR000159; RA.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   Pfam; PF00560; LRR; 16.
DR   Pfam; PF00481; PP2C; 2.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00044; CYCc; 1.
DR   SMART; SM00370; LRR; 7.
DR   SMART; SM00369; LRR_TYP; 2.
DR   SMART; SM00332; PP2Cc; 1.
DR   SMART; SM00331; PP2C_SIG; 1.
DR   SMART; SM00314; RA; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
SQ   SEQUENCE   1690 AA;  189400 MW;  8AB32D9B8C747AB0 CRC64;


  Query Match             31.1%;  Score 50;  DB 3;  Length 1690;
  Best Local Similarity   39.1%;  Pred. No. 2.9e+02;
  Matches    9;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFYFKIDNVK 25
           |: | |||  || : |  :| ::
Db     442 TYEIEKLNVANNPSIYLPLDFIQ 464
```

Search completed: July 1, 2002, 16:29:47
Job time: 701 sec